DOCUMENT RESUME

ED 390 391 IR 055 733

AUTHOR	Drabenstott, Karen M.; Weller, Marjorie S.
TITLE	Testing a New Design for Subject Access to Online Catalogs.
INSTITUTION	Michigan Univ., Ann Arbor. School of Information and Library Studies.
PUB DATE	Aug 95
NOTE	390p.; For a related article, see EJ 478 035.
AVAILABLE FROM	School of Information and Library Studies, University of Michigan, 304 West Engineering Building, 550 East University Ave., Ann Arbor, MI 48109-1092 (in U.S.: $18 prepaid; in Canada: $21 prepaid; outside U.S. and Canada: $44 prepaid; make checks payable to University of Michigan).
PUB TYPE	Statistical Data (110) -- Reports - Descriptive (141) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC16 Plus Postage.
DESCRIPTORS	*Access to Information; Data Analysis; Design Requirements; Indexing; *Information Retrieval; Information Seeking; *Library Catalogs; *Online Catalogs; Research and Development; *Search Strategies; *Subject Index Terms; User Needs (Information); User Satisfaction (Information); Use Studies
IDENTIFIERS	Query Formulations; Search Behavior; *Search Trees; Transaction Log Analysis

ABSTRACT

Over the last decade, studies of online catalogs have revealed that they frequently discourage users. Subject queries often fail to produce retrievals or produce retrieval sets that are too large and unwieldy to be easily scanned. Research shows that users are seeking alternative approaches to those that manipulate the subject headings in catalog records. This document reports on the development and evaluation of the ASTUTE experimental online catalog, which makes use of search trees to enhance subject access. The project investigated whether search trees actually improved the search performance of subject searchers at the terminal and whether users liked the concept of controlled system responses; they also were designed to assemble a list of characteristics of the types of queries that ASTUTE failed to answer. Chapter 1 describes circumstances and factors that gave rise to the ASTUTE project. Chapter 2 reviews previous research on subject queries and search trees. Chapters 3, 4, and 5 describe the development of the experimental catalog. Chapter 6 discusses transaction logs and related methodologies that were used in testing the system's retrieval. Chapters 7 through 12 present analysis of the online retrieval test data, including user behavior, precision and recall, and failure analysis. They also explore resulting ideas for redesigning the search trees. Chapter 13 is a summary of project activities, findings, and recommendations. Some specifics on search trees, sample questionnaires and consent forms, and transaction log records are appended. (Contains 99 tables and 113 figures.) (BEW)

ED 390 391

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Testing a New Design for Subject Access to Online Catalogs

by
Karen M. Drabenstott and Marjorie S. Weller
with the research assistance of Jeffrey M. Holden

School of Information and Library Studies
The University of Michigan
304 West Engineering Building
550 East University Avenue
Ann Arbor, Michigan 48109-1092 USA

August 1995

IR055733

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Karen M. Drabenstott
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Testing a New Design for Subject Access to Online Catalogs

by
Karen M. Drabenstott and Marjorie S. Weller
with the research assistance of Jeffrey M. Holden

School of Information and Library Studies
The University of Michigan
304 West Engineering Building
550 East University Avenue
Ann Arbor, Michigan 48109-1092 USA

August 1995

Table of Contents

Table of Contents i

List of Tables v

List of Figures viii

Acknowledgments xii

Obtaining Copies of this Report xiv

About the Authors xvi

1 Project Objectives and Research Questions 1

2 Literature Review 7
- 2.1 Studies of the Subject Terms Users Enter into Online Catalogs 7
- 2.2 Subject Searching Improvements 15
- 2.3 Subject Searching Improvements to Provide Useful Information 16
- 2.4 Subject Searching Functionality and Search Trees 19
- 2.5 Chapter Summary 20

3 Machine-readable Data for 23
- 3.1 Introduction 23
- 3.2 Computer Equipment 23
- 3.3 Machine-readable Bibliographic Data 24
- 3.4 Subject Authority Data 29
- 3.5 Chapter Summary 31

4 Search Trees for Subject Searching 33
- 4.1 Introduction 33
- 4.2 Search Tree Design 33
- 4.3 Initial Search Tree 34

4.4 Search Trees for Subject Queries Generally 35
4.5 Search Tree for Personal-name Queries 44
4.6 Chapter Summary 45

5 ASTUTE Subject Searching Functionality 47
5.1 Introduction 47
5.2 FoxPro Database Management System 47
5.3 Exact approach to Subject Searching 48
5.4 Alphabetical Approach to Subject Searching 58
5.5 Keyword-in-main-heading Search 63
5.6 Keyword-in-subdivided-heading Search 71
5.7 Title-keyword Search 75
5.8 Remaining Keyword Searches 79
5.9 Personal-name Searches 84
5.10 Pinstripe System's Random Selection Algorithm 94
5.11 Displaying Bibliographic Records 95
5.12 Chapter Summary 97

6 Methods Used to Test the New Subject Access Design 99
6.1 Introduction 99
6.2 Previous Studies Comparing Systems 100
6.3 Online Questionnaire Administration in Catalog Use Studies 102
6.4 Comparison Search Experiment (Library Patrons) 103
6.5 Comparison Search Experiment (Library Staff) 113
6.6 Recording Search Experiments to a Transaction Log 115
6.7 Timeout Function 115
6.8 Design of the Comparison Search Experiment 115
6.9 Chapter Summary 120

7 Comparison Search Experiment Participants 122
7.1 Introduction 122
7.2 Michigan Pretest 122
7.3 Installation Period Length 123
7.4 First Comparison Search Experiment Administration at UM-Dearborn 125
7.5 Second Comparison Search Experiment Administration at UM-Dearborn 127

7.6 Comparison Search Experiment Administration at Earlham College 139
7.7 Reliability of Relevance Assessments 148
7.8 Reliability of Post-search Questionnaire Responses 150
7.9 Chapter Summary 151

8 Characteristics of System Users and Searches 154
8.1 Introduction 154
8.2 Experimental System Users 154
8.3 Time Spent Searching 157
8.4 User Queries 160
8.5 Retrieved and Displayed Titles 163
8.6 Chapter Summary 171

9 Precision and Estimated Recall in the Blue and Pinstripe Systems 173
9.1 Introduction 173
9.2 Precision in Experimental Catalog Searches 173
9.3 Estimated Recall Scores 184
9.4 Post-search Questionnaire Responses 187
9.5 Chapter Summary 199

10 Failure Analysis of User Searches 202
10.1 Introduction 202
10.2 Successful and Unsuccessful Searches at UM-Dearborn 202
10.3 Successful and Unsuccessful Searches at Earlham 232
10.4 Chapter Summary 252

11 Failure Analysis of Subject Searching Capabilities 255
11.1 Introduction 255
11.2 Controlled Vocabulary Searching Based on the Exact Approach 257
11.3 Keyword-in-Subdivided-Heading Search 275
11.4 Title-keyword Searches 283
11.5 Keyword in Subject Heading Field Searches 285
11.6 Keyword-in-record Searches 286
11.7 Subject Searches for Personal Names 290
11.8 Librarians' Subject Searching Experiences 299
11.9 Chapter Summary 304

12 Redesigning Search Trees for Responsive Subject Searching 307
12.1 Introduction 307
12.2 The Need for Automatic Spelling Correction 307
12.3 Vocabulary Problems and Search Trees 312
12.4 Redesigned Search Trees 316
12.5 Chapter Summary 328

13 Highlights of Project Activities 331
13.1 Project Overview 331
13.2 Development of the ASTUTE Experimental Online Catalog 332
13.3 Methods Used to Test the New Subject Access Design 334
13.4 Comparison Search Experiment Participants 335
13.5 Characteristics of System Users and Searches 337
13.6 Precision and Estimated Recall in the Blue and Pinstripe Systems 338
13.7 Post-search Questionnaire Responses Regarding System Performance 340
13.8 Failure Analysis of User Searches 340
13.9 Failure Analysis of Subject Searching Approaches 343
13.10 Redesigning Search Trees for Responsive Subject Searching 345
13.11 A New Design for Subject Access to Online Catalogs 346

Appendix A. Form Subdivisions 348

Appendix B. Control-field Elements 350

Appendix C. End-user Questionnaire 352

Appendix D. Cover Letter and Consent Form 355

Appendix E. Post-search Questionnaire Administered to Library Staff 357

Appendix F. ASTUTE Transaction Log Records 360
F.1 Level 1 Transaction Log Records 360
F.2 Level 2 Transaction Log Records 362
F.3 Level 3 Transaction Log Records 367
F.4 Level 4 Transaction Log Records 370

List of Tables

3.1 Authority Records Selected 30

4.1 Sequence of personal-name query elements 44

5.1 Fields/Subfields in Exact Search Approach Database 50
5.2 Exact Search Approach Options 51
5.3 Alphabetical Search Approach Options 59
5.4 Keyword-in-main-heading Search Approach Options 66
5.5 Fields/Subfields in the Keyword-in-subdivided-heading Search Database 72
5.6 Keyword-in-subdivided-heading Search Approach Options 72
5.7 Fields/Subfields in the Title-keyword Search Database 76
5.8 Title-keyword Search Options 77
5.9 Fields/Subfields in the Keyword-in-record database 80
5.10 Subject Heading Keyword and Keyword-in-record Search Options 82
5.11 Fields/Subfields in the Personal-name, Keyword-in-record Database ..87
5.12 Personal-name, Keyword-in-record Search Options 88
5.13 Alphabetical Search Approach Options for Personal -name Queries 92
5.14 Fields/Subfields for Bibliographic Record Displays 95

6.1 Variables Measured in Comparison Search Experiment 116
6.2 Comparison Search Experiment Procedures Summary 117
6.3 Full and Partial Administrations of the Comparison Search Experiment 117

7.1 System Installation Information 125
7.2 Usable Administrations in the Second Comparison Search Experiment at UM-D 130
7.3 Unusable Administrations of the Second Comparison Search Experiment at UM-D 131
7.4 Usable Comparison Search Experiment Administrations at Earlham 142
7.5 Unusable Comparison Search Experiment Administrations at Earlham 144

8.1 Major Field of Study 156

8.2 Time Spent Per Search Administration Type....157
8.3 Time Spent Searching Each System....159
8.4 Characteristics of User Queries....161
8.5 T-tests Comparing Query Lengths....162
8.6 "Opportunities" by Search Administration Type....164
8.7 Retrievals by Search Administration Type....165
8.8 Retrieved and Displayed Titles....166
8.9 Titles Displayed Before and After Resolution of Conflicting Relevance Assessments....168
8.10 Comparing Relevance Assessments Before and After Resolution of Conflicts....170

9.1 T-test Results for Precision (UM-D)....176
9.2 T-test Results for Precision (Earlham)....177
9.3 T-test Results for Precision in the Blue System (UM-D)....179
9.4 T-test Results for Precision in the Pinstripe System (UM-D)....181
9.5 T-test Results for Precision in the Blue System (Earlham)....183
9.6 T-test Results for Precision in the Pinstripe System (Earlham)....183
9.7 Topics to Extend to ASTUTE Features....199

10.1 Successful and Unsuccessful Searches at UM-D....203
10.2 User Perseverance....204
10.3 Specificity of User Queries....206
10.4 Database Failure....208
10.5 Large Retrievals....209
10.6 Vocabulary Problems....211
10.7 Subject Searching Approach Failures in the Pinstripe System....212
10.8 Subject Searching Approach Failures in the Blue System....213
10.9 Navigational Problems....218
10.10 Reclassified Searches (Both systems, UM-D)....219
10.11 User Perseverance....222
10.12 Specificity of User Queries....223
10.13 Large Retrievals....224
10.14 Vocabulary Problems....225
10.15 Subject Searching Approach Failures in the Pinstripe System....226
10.16 Reclassified Searches (Single system, UM-D)....229
10.17 Reclassified Successful and Unsuccessful Searches at UM-D....231
10.18 Successful and Unsuccessful Searches at Earlham....233
10.19 User Perseverance....234
10.20 Specificity of User Queries....236
10.21 Database Failure....237
10.22 Vocabulary Problems....239

10.23 Subject Searching Approach Failures in the Pinstripe System 240
10.24 Reclassified Searches (Both systems, Earlham) 241
10.25 User Perseverance 244
10.26 Specificity of User Queries 245
10.27 Subject Searching Approach Failures in the Pinstripe System 247
10.28 Subject Searching Approach Failures in the Blue System 247
10.29 Navigational Problems 248
10.30 Reclassified Searches (Single system, Earlham) 249
10.31 Reclassified Successful and Unsuccessful Searches at Earlham 251

11.1 Matches in Blue-system Searches 255
11.2 Queries Submitted to the Pinstripe System 257
11.3 Browsing/Display Opportunities (UM-D) 260
11.4 Browsing/Display Opportunities (Earlham) 261
11.5 Browsing in Successful Controlled Vocabulary Searches 264
11.6 Simple, Successful Controlled Vocabulary Searches 265
11.7 Successful Controlled Vocabulary Searches With a Little Browsing ..266
11.8 Successful Controlled Vocabulary Searches With More than a Little Browsing 267
11.9 Successful Controlled Vocabulary Searches With Expanding 268
11.10 Controlled Vocabulary Searches Marred by User Perseverance 270
11.11 Navigation Problems in Controlled Vocabulary Searches 273
11.12 Queries and Matching Terms in Pinstripe Alphabetical Searches 274
11.13 Keyword-in-subdivided-heading Searches 276
11.14 Successful Keyword-in-subdivided-heading Searches 279
11.15 Blue-system Search Outcomes for Failed Keyword-in-subdivided-heading Searches 280
11.16 Perseverance in Keyword-in-subdivided-heading Searches 282
11.17 Title-keyword Searches 284
11.18 Frequently-occurring Subject Headings in Title-keyword Searches...285
11.19 Keyword-in-record Searches 288
11.20 Frequently-occurring Subject Headings in Keyword-in-record Searches 289
11.21 Subject Searches for Personal Names 297

12.1 User Actions Following System Message Regarding Possible Spelling Errors 309
12.2 Succeeding Queries on Different Topics 310
12.3 Misspelled Queries 310
12.4 Queries with Words Added or Deleted 311
12.5 Stemming Queries to Enhance Retrieval 314
12.6 Queries Suited to the Best-match Approach 315

List of Figures

1.1. Project activities ... 4

3.1 Title bearing terms for control-field codes ... 29

4.1 Initial search tree ... 34
4.2A Search tree for the exact approach ... 36
4.2B Search tree for the exact approach (contd.) ... 37
4.3 Search tree for one-word queries ... 38
4.4 Search tree for multi-word queries featuring the alphabetical approach ... 40
4.5A Search tree for multi-word queries featuring keyword approaches ... 41
4.5B Search tree for multi-word queries featuring keyword approaches (contd.) ... 43

5.1 User entry of "civil rights" query ... 52
5.2 Intermediary matched terms list for "civil rights" ... 53
5.3 Exact approach main menu for "Civil rights" ... 54
5.4 Scope note for "Civil rights" ... 54
5.5 Bibliographic record bearing subject heading "Civil rights" ... 55
5.6 Broader terms for "Civil rights" ... 56
5.7 New exact search main menu for "Discrimination" ... 56
5.8 Place subtopics under "Civil rights" ... 57
5.9 User entry of "computer crime" query ... 61
5.10 Alphabetical list of subject headings provoked by "computer crime" query ... 61
5.11 User selection of "Computer crimes" ... 62
5.12 Exact approach main menu for "Computer crime" ... 62
5.13 Bibliographic record bearing subject heading "Computer crimes" ... 63
5.14 User entry of "control systems" query ... 67
5.15 Alphabetical keyword list for "trade and industry" query ... 68
5.16 Selection of a listed heading from the alphabetical keyword list ... 69
5.17 Exact approach main menu for "Rubber industry and trade" ... 69
5.18 Selection of "Information services" subtopic in an exact search ... 70

5.19 Bibliographic record bearing subdivided heading "Rubber industry and trade" 70
5.20 User entry of "women in history" query 73
5.21 Alphabetical keyword list for "women in history" query 74
5.22 User selection of subdivided heading 74
5.23 Bibliographic record bearing selected subdivided heading 75
5.24 User entry of "black soldiers" query 78
5.25 Bibliographic record bearing titles words matching user query 78
5.26 Bibliographic record bearing title words matching user query 79
5.27 User entry of "electric powered automobiles" query 83
5.28 Bibliographic record bearing words matching user query 83
5.29 User entry of last name query element 89
5.30 User entry of first name query element 89
5.31 User entry of topic query element 90
5.32 Bibliographic record bearing words in user query 90
5.33 User entry of personal-name query for "tecumseh" 92
5.34 Alphabetical list of personal-name subject headings in response to "tecumseh" query 93
5.35 User selection of personal-name subject heading 93
5.36 Bibliographic record bearing the subject heading "Tecumseh" 94
5.37 Prompting users for relevance assessments 97

6.1 First introductory screen 104
6.2 Second introductory screen 105
6.3 Third introductory screen 106
6.4 Fourth introductory screen 106
6.5 Fifth introductory screen 107
6.6 Pop-up window with question on test participation 108
6.7 First-pre-search question on previous ASTUTE use 109
6.8 Second pre-search question on using other computer systems 109
6.9 Third-pre-search question on major field of study 110
6.10 Question on personal names in subject queries 111
6.11 Relevance assessment categories 112

7.1 Usable and unusable administrations of the first Comparison Search Experiment at UM-Dearborn 126
7.2 Usable and unusable administrations of the second Comparison Search Experiment at UM-Dearborn 127
7.3 Search administrations at UM-D on a daily basis 128
7.4 Usable search administrations at UM-D by day of the week 129
7.5 Usable and unusable Comparison Search administrations at Earlham140
7.6 Search administrations at Earlham on a daily basis 141
7.7 Usable search administrations at Earlham by day of the week 142

8.1 Previous use of ASTUTE.......155
8.2 Other computer use.......155
8.3 Time spent searching per event at UM-D.......158
8.4 Time spent searching per event at Earlham.......159

9.1 Precision by search administration type (UM-D).......174
9.2 Precision by search administration type ((Earlham).......175
9.3 Comparing precision in Blue-system controlled vocabulary and free-text searches (UM-D).......178
9.4 Comparing precision in Pinstripe-system controlled vocabulary and free-text searches (UM-D).......180
9.5 Comparing precision in Blue-system controlled vocabulary and free-text searches (Earlham).......182
9.6 Estimated recall (UM-D).......185
9.7 Estimated recall (Earlham).......186
9.8 Comparing the number of useful titles retrieved.......188
9.9 Satisfaction with the search.......189
9.10 System preference.......190
9.11 Ease of giving instructions to the experimental systems.......191
9.12 Ease of getting instructions from the experimental systems.......192
9.13 Clarity of the experimental systems.......193
9.14 Efficiency of the experimental systems.......194
9.15 Helpfulness of giving new ideas for subject searching.......195
9.16 Importance of system features.......196
9.17 User familiarity with subject searched.......197
9.18 Extending ASTUTE features to other topics.......198

10.1 Exact-search options for "Historic buildings".......215
10.2 Exact-search options for "Historic buildings — Michigan".......215
10.3 Exact-search options for "Historic buildings — Michigan — Detroit".......216
10.4 Exact-search options for "Historic buildings — Michigan — Detroit — Conservation and Restoration".......216
10.5 Reclassified unsuccessful searches (both systems, UM-D).......220
10.6 System response to misspelled query words.......228
10.7 Reclassified searches (single system, UM-D).......230
10.8 Reclassified searches (both systems, Earlham).......242
10.9 Reclassified searches (singe system, Earlham).......249

11.1 Precision in controlled vocabulary searches.......258
11.2 Number of opportunities per search (UM-D and Earlham).......262
11.3 Exact-approach main menu.......271
11.4 Precision in keyword-in-subdivided-heading searches.......277
11.5 Precision in keyword-in-record searches.......287

11.6 Asking users whether their queries involve names....290
11.7 Precision in personal-name subject searches....296

12.1 Informing users of possible misspellings....308
12.2 Checking queries from left to right for possible misspellings....308
12.3 Redesigned initial search tree....317
12.4A Search tree for the exact approach....318
12.4B Search tree for the exact approach (contd.)....319
12.5A Redesigned search tree for one-word queries....320
12.5B Redesigned search tree for one-word queries (contd.)....321
12.6A Redesigned search tree for multi-word queries....324
12.6B Redesigned search tree for multi-word queries (contd.)....325
12.6C Redesigned search tree for multi-word queries (contd.)....327

Acknowledgments

The ASTUTE project team consisted of Karen Markey Drabenstott, Marjorie S. Weller, and Jeffrey M. Holden.

In this report, we describe a test of a new design for subject access to online catalogs. The Department of Education provided support for the test. We are especially grateful to Neal K. Kaske, Senior Associate, Office of Library Programs, Department of Education, who encouraged our research efforts and answered questions that arose during the project.

There were many staff members at participating libraries who contributed to the success of this Department of Education-sponsored research project. Timothy F. Richards, Head, Mardigian Library, University of Michigan-Dearborn (UM-D), and Evan Ira Farber, Director, Lilly Library, Earlham College, welcomed us to their libraries to collect bibliographic data and conduct search experiments with their libraries' patrons and staff. Robert Kelly at UM-D and Michael Bowden at Earlham helped us install the experimental online catalog, monitored the system during its installation at their libraries, transferred or sent transaction logs to us electronically or via ground mail, recruited library staff to participate in the search experiment, and answered our many questions about their libraries' bibliographic records and users. When we realized our first data collection effort at UM-D had been unsuccessful, both Tim and Bob invited us back to Mardigian Library to conduct a second search experiment. When we experienced difficulty in the speed of downloading bibliographic records from Earlham's online catalog database to floppy disk, Michael loaned us his new, speedy microcomputer for several days for downloading tasks. We also thank library staff members at UM-D and Earlham for taking part in the search experiment.

Advice from our project consultants was encouraging and supportive throughout the project. We appreciate the efforts of Martin Dillon and Paul Kantor who advised us on the experimental design. Paul Kantor reviewed specifications for the transaction logging capability and made suggestions on analyzing logged data. William H.

Mischo and Paul Kantor reviewed system functionality, drafts of the final report, and offered helpful suggestions on each.

At the Library of Congress, Susan Tarr of Cataloging Distribution Service made it possible for us to obtain a complimentary copy of CD/MARC Subjects which we put to work in the experimental online catalog.

At the University of Michigan-Ann Arbor, we appreciate the assistance of Michael G. Moore, presently Systems Development Coordinator, Population Studies Center, who answered technical questions about FoxPro, helped us download database tapes from UM-D, and transferred the ASTUTE development application into a user-executable application. We thank Jeffrey M. Holden who submitted transaction log data to statistical tests for the statistical and failure analyses. The authors are also grateful to Robert Royce who transferred and converted screen images from FoxPro to Microsoft Word and provided expert assistance in the design and layout of this report.

Obtaining Copies of this Report

Testing a New Design for Subject Access to Online Catalogs may be obtained over the Internet using anonymous FTP. The FTP host name is "sils.umich.edu." When the system asks you for a user name and password, type the user name "anonymous" and type your electronic mail address (or last name) for the password. The directory containing this report is /pub/papers/DoED.

The files for *Testing a New Design* are unusually large because they contain many images showing screens in the ASTUTE experimental online catalog. Files bearing the ".ps" suffix in the anonymous FTP directory are files in Postscript format. We strongly suggest you obtain this report in the Postscript file format to preserve layout, paging, figures, tables, and type fonts. The sizes and contents of Postscript files in the anonymous FTP directory are as follows:

File name	Contents	Size
DoED1.ps	Introductory material and chapters 1–4, pages i–xvi, 1–46	2.2 MB
DoED2a.ps	Chapter 5, pages 47–57	10.3 MB
DoED2b.ps	Chapter 5, pages 58–70	14.1 MB
DoED2c.ps	Chapter 5, pages 71–83	11.6 MB
DoED2d.ps	Chapter 5, pages 84–98	11.6 MB
DoED3.ps	Chapter 6, pages 99–121	14 MB
DoED4.ps	Chapters 7–10, pages 122–254	7.9 MB
DoED5.ps	Chapters 11–13, appendixes A–F, pages 255–370	6.3 MB

Testing a New Design is also available as a self-extracting Stuffit archive. The name of this archive is "DoED.sea.hqx." To retrieve it, use the FTP host name "sils.umich.edu." When the system asks you for a user name and password, type the user name "anonymous" and type your electronic mail address (or last name) for the password. The directory containing this archive is

/pub/papers/DoED. You will also need a copy of the BinHex utility for the Macintosh. This is a free, public domain utility which can be obtained via anonymous FTP from "archive.umich.edu" in the directory /mac/utilities/compression. Use Binhex to decode the file. To open the self-extracting archive, double click on the archive's icon. You will need Microsoft Word 5.1a for the Macintosh to display the eight word-processing files in the archive. The names of these eight files correspond to the eight Postscript files listed in the table above. Unstuffed, they will range in size from 270 thousand bytes to 1.1 megabytes.

Published (bound and printed) copies of *Testing a New Design* may be obtained from the School of Information and Library Studies at the following rates: (1) $18 prepaid, 1st class delivery in the United States, (2) $21 prepaid, first class delivery in Canada, or (3) $44 prepaid, air mail delivery outside the United States and Canada. Please make checks payable to the University of Michigan and send them to the School of Information and Library Studies, University of Michigan, 304 West Engineering Building, 550 East University Avenue, Ann Arbor, Michigan 48109–1092 USA. The authors have submitted the report to ERIC, and, pending acceptance and processing, will be available in microfiche and hardcopy formats from ERIC.

About the Authors

Karen Markey Drabenstott is an Associate Professor in the School of Information Studies at the University of Michigan. The impetus for the research discussed in this study were findings from a Council on Library Resources-sponsored research project which are given in the book entitled *Using Subject Headings for Online Retrieval: Theory, Practice, and Potential* written by Karen Drabenstott and Diane Vizine-Goetz and published by Academic Press in 1994. Support from the Council also enabled Karen to research and write the *Analytical review of the library of the future* in 1994. This analytical bibliography and synthesis of published literature on the library of the future prepared her for her current role as faculty coordinator of the Kellogg Coalition on Information Science, Technology, and Library Education (KRISTaL-Ed), a five-year, multi-million dollar project supported by the Kellogg Foundation to provide national leadership in educating information professionals for the 21st century.

Karen joined the faculty of The University of Michigan in January 1987. From 1981 to 1986, she was a research scientist in the Office of Research at OCLC. She received her B.A. from The Johns Hopkins University and her M.L.S. and Ph.D. from the School of Information Studies at Syracuse University.

Marjorie S. Weller designed, developed, and implemented the ASTUTE experimental online catalog using the FoxPro database management system. She is currently a Programmer Analyst I in the Medical Center Information Technologies where she programs financial information systems for the Medical School of the University of Michigan. She received an A.S. in Computer Science from Henry Ford Community College.

Jeffrey M. Holden generated the statistics for the quantitative analysis of online retrieval test data. He holds an A.B. and M.B.A. from Michigan State University and is pursuing a master's degree in the School of Information and Library Studies.

1 Project Objectives and Research Questions

The 1980s began with a handful of libraries offering end users direct computerized access to library holdings through the online catalog component of their local online system. Ensuing the introduction of online catalogs in libraries were studies of online catalog users. One of the earliest and most comprehensive studies was a nationwide survey of online catalog use in twenty-nine libraries sponsored by the Council on Library Resources (CLR) (Matthews et al. 1983). The remark "Sacred cows are being strewn all over the landscape" characterized survey findings because they challenged traditional beliefs about catalog users and uses (Besant 1982, 160). In particular, survey findings about subject access — the predominance of subject searching, the many users who experienced difficulties conducting subject searches, and the large number of suggestions from users for improvements to subject searching — were completely unexpected.

Nationwide survey findings stimulated a decade of research in the area of subject access. CLR, OCLC (Online Computer Library Center), BLRDD (British Library Research and Development Department), and the (U.S.) Department of Education have been exemplary in their support of research on subject access (Drabenstott 1991, 64–6). Library staff also called for improvements to online catalogs to reflect user needs. Since staff play a large role in system selection, system designers have been especially receptive to staff demands for improved ease of use, enhancements to search functionality (i.e., keyword, Boolean-based approaches), and extensions of online catalog functionality to other databases (Hildreth 1991, 21).

Online catalogs are now commonplace in academic and public libraries. Telecommunications make it possible for end users to access a library's holdings without having to set foot in the library. National and international computer networks enable users to access online catalogs of libraries throughout the United States and the world.

Despite a decade of online catalog research, development, and deployment, many of the same problems that the earliest online catalog searchers experienced plague today's users, particularly in the area of subject access. Examples are:

- One-third of the subject queries users enter into online catalogs fail to produce retrievals. (Lynch 1989, 52)
- When searches produce retrievals, large retrievals discourage users from scanning results. (Markey and Demeyer 1986, 277; Van Pulis and Ludy 1988, 528; Lynch 1989, 52)
- The few instances when users are successful matching the catalog's controlled vocabulary are when they enter one-word queries for topics or places. (Van Pulis and Ludy 1988, 527; Carlyle 1989, 44; Drabenstott and Vizine-Goetz, 1994, 168)
- Users have become so discouraged with the results of subject searches that they are seeking alternative approaches to those that manipulate the subject headings in cataloging records. (Larson 1991)

These problems are indicative of the need for a new design for subject access to online catalogs. The foundation for the new design are findings from an empirical study of the subject terms users enter into online catalogs (Drabenstott and Vizine-Goetz 1994). The new design requires that online catalogs have a wide range of subject searching capabilities and search trees to govern the system's selection of searching capabilities in response to user queries. Search trees hold much promise for assuming the burden of determining which subject searching approach is likely to produce useful information in response to user queries.

This report describes a research project that tested the new subject access design. Project objectives were to:

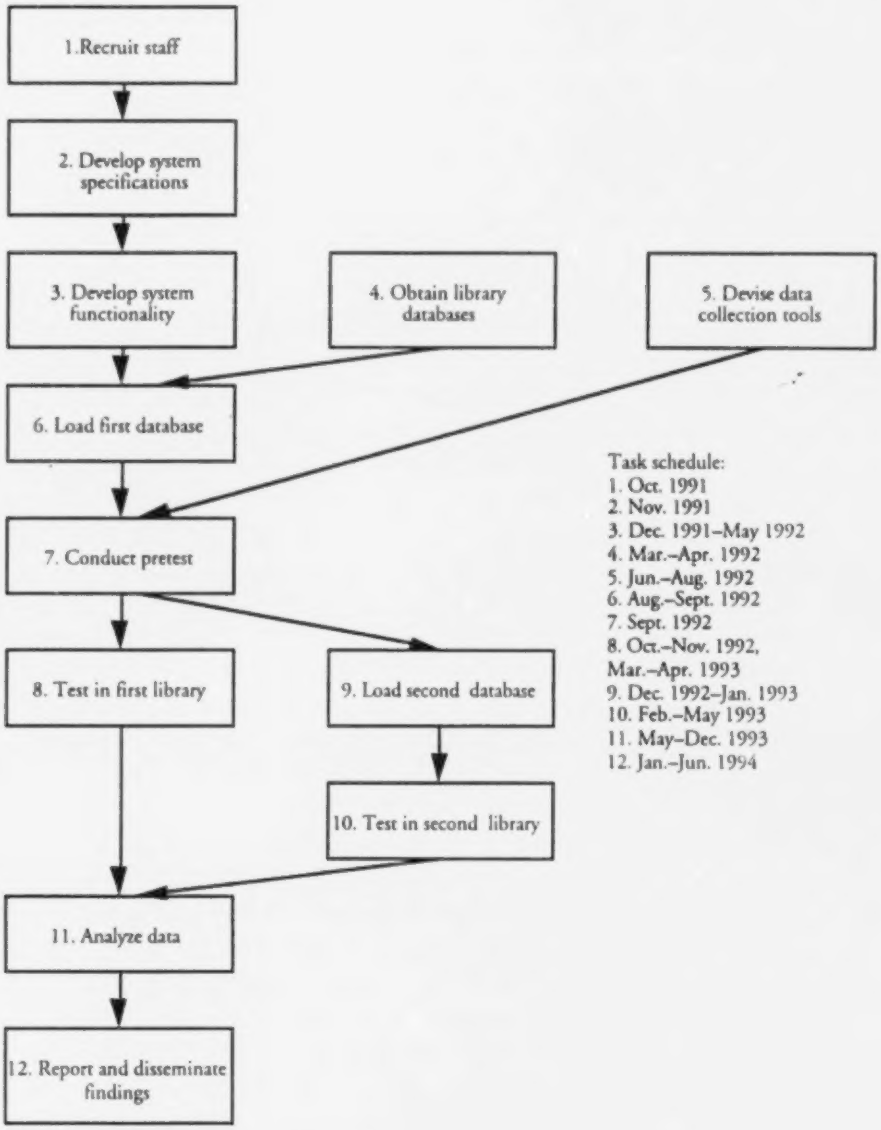
1. Demonstrate subject searching in an experimental online catalog with a wide range of subject searching capabilities and with search trees to govern system selection of a particular capability in response to user queries.
2. Test the retrieval effectiveness of the experimental online catalog with search trees by comparing its performance with the performance of an experimental online catalog in which subject searching approaches are assigned at random.
3. Evaluate the demonstration and test results of retrieval effectiveness and disseminate the research findings through publications in the professional literature.

Underlying the design of the ASTUTE experimental online catalog were two important assumptions: (1) subject searching functionality was limited to searching capabilities in operational online catalogs, and (2) titles and subject headings (based on the Library of Congress Subject Heading system) were the principal indicators of subject content in machine-readable cataloging (MARC) records.

The results of the research project have the potential for generating a new design. Since ASTUTE's subject searching functionality was limited to capabilities in operational catalogs, system designers would not have to discard their operational online catalogs because the new design would be based on subject searching approaches that are already available in existing systems. Similarly, the subject content of ASTUTE's cataloging databases was based on the same information in the cataloging databases of most American libraries. Thus, library technical services staff would not have to discard their existing databases or introduce new methods to take advantage of the new design. The new design would, however, require system designers to add a wider range of subject searching functionality than their systems currently support and search trees to control system responses.

The research project was composed of twelve activities grouped into three phases: (1) system development (tasks 1–6, 9), (2) system testing (tasks 7, 8, 10), and (3) data analysis and reporting (tasks 11–12). Figure 1.1 shows a flowchart of project tasks and gives a schedule of project tasks.

Seven activities (1–6, 9) were required in the system-development phase. This phase featured the development of an experimental online catalog with a wide range of subject searching functionality and search trees to govern the system's selection of a subject searching approach in response to user queries. Bibliographic records from the online catalogs of Mardigian Library at the University of Michigan-Dearborn (UM-D) and Lilly Library at Earlham College were loaded into separate, searchable databases of the experimental online catalog. Three activities (7–8, 10) were required in the system testing phase. Data collection procedures and instruments were developed, pretested, revised, and administered in online retrieval experiments with library patrons and staff at the libraries of UM-D and Earlham College to answer the following three research questions:


1.Recruit staff
2. Develop system specifications
3. Develop system functionality
4. Obtain library databases
5. Devise data collection tools
6. Load first database
7. Conduct pretest
8. Test in first library
9. Load second database
10. Test in second library
11. Analyze data
12. Report and disseminate findings

Task schedule:
1. Oct. 1991
2. Nov. 1991
3. Dec. 1991–May 1992
4. Mar.–Apr. 1992
5. Jun.–Aug. 1992
6. Aug.–Sept. 1992
7. Sept. 1992
8. Oct.–Nov. 1992,
Mar.–Apr. 1993
9. Dec. 1992–Jan. 1993
10. Feb.–May 1993
11. May–Dec. 1993
12. Jan.–Jun. 1994

Figure 1.1. Project activities

1. Do search trees improve the search performance of subject searchers at the online catalog?
2. Do subject searchers prefer an online catalog that controls system responses and searching approaches to an online catalog without such controls?

3. What are the characteristics of queries that cannot be answered using the ASTUTE experimental online catalog?

Two activities (11–12) made up the evaluation phase of the project. Collected data were described, analyzed, interpreted, and synthesized in a final report to the Department of Education. The ASTUTE Project Team also plans to disseminate findings in journals and conference proceedings to the library and information science communities.

This report of the ASTUTE Project consists of thirteen chapters. Chapter 1 describes factors that led to the project director's research proposal on testing a new subject access design to online catalogs and the research questions and objectives of the project. Chapter 2 reviews previous research on search trees, subject searching functionality, and user queries for subjects. Chapters 3–5 describe the development of the experimental online catalog, in particular, its bibliographic and authority databases, subject searching functionality, and search trees. Chapter 6 focuses on the methods used in online retrieval tests and includes a discussion of the catalog's transaction logging capability that was used to record users' subject searches, relevance assessments, and administer pre- and post-search questionnaires. The results of the analysis of online retrieval test data are covered in chapters 7–12. Results chapters include recommendations on reconfiguring and enhancing search trees to accommodate the wide range of subject queries users enter into online catalogs. Chapter 13 is a summary of project activities, findings, and recommendations.

The data and analyses of this research project generally increase our understanding of the subject terms users enter into online catalogs. The project builds on a new design for subject access to online catalogs that enlists search trees to control system responses and determine appropriate subject searching approaches to user queries. The project will ensure that future subject searching implementations in online catalogs are responsive to the wide variety of user queries for subjects.

References

Besant, Larry. 1982. "Early survey findings: users of public online catalogs want sophisticated subject access." *American Libraries* 107, 19 (March): 160.

Carlyle, Allyson. 1989. "Matching LCSH and user vocabulary in the library catalog." *Cataloging & Classification Quarterly* 10, (1/2): 37–63.

Drabenstott, Karen M. 1991. "Online Catalog User Needs and Behaviors." In *Think Tank on the Present and Future of the Online Catalog: Proceedings*, edited by Noelle Van Pulis, 59–83. Chicago: Reference and Adult Services Division, American Library Association. RASD Occasional Papers, 9.

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Drabenstott, Karen M., and Diane Vizine-Goetz. 1990. "Search trees for subject searching in online catalogs." *Library Hi Tech* 8, 3: 7–20.

Hildreth, Charles R. 1991. "Advancing toward the E[3] OPAC: the imperative and the path." In *Think Tank on the Present and Future of the Online Catalog: Proceedings*, edited by Noelle Van Pulis, 17–38. Chicago: Reference and Adult Services Division, American Library Association. RASD Occasional Papers, 9.

Larson, Ray R. 1991. "The Decline of Subject Searching: Long-term Trends and Patterns of Index Use in an Online Catalog." *Journal of the American Society for Information Science* 42, 3 (April): 197–215.

Lynch, Clifford A. 1989. "Large Database and Multiple Database Problems in Online Catalogs." In *OPACs and beyond.* Dublin, Ohio: OCLC.

Markey, Karen, and Anh Demeyer. 1986. *Dewey Decimal Classification Online Project: Final Report to the Council on Library Resources.* Dublin, Ohio: OCLC. OCLC Research Report OCLC/OPR/RR–86/1.

Matthews, Joseph R., et al. 1983. *Using Online Catalogs: A Nationwide survey.* New York: Neal-Schuman.

Van Pulis, Noelle, and Lorene Ludy. 1988. "Subject Searching in an Online Catalog with Authority Control." *College & Research Libraries* 49, 6 (November): 523–33.

2 Literature Review

2.1 Studies of the Subject Terms Users Enter into Online Catalogs

The computer's ability to record every system response and user action input into the online catalog provides researchers with a very accurate tool for collecting the subject terms users enter into online catalogs. Such user-system interaction is called a transaction log.

At the present time, standards do not exist for transaction logs. Consequently, the content and format of transaction logs varies from system to system. Computer programs written to analyze transaction log data from one system must be rewritten to analyze data from a different system.

An important advantage of transaction logs is the unobtrusiveness of this data collection approach. A computer program collects user-system interaction data unbeknownst to catalog users. Consequently, users are not likely to alter their catalog-seeking behavior because they do not know that the terms they are entering into the catalog are being recorded for analysis. Another advantage of transaction logging is the ability to collect voluminous amounts of data in machine-readable form for subsequent analyses using computer programs.

Transaction logs also have disadvantages. Logged queries do not always provide researchers with an exact representation of what users are looking for. This information could be obtained by supplementing transaction logs with observations of user searches, personal interviews, or a combination of the two. Logs also do not provide researchers with accurate demarcations of searches entered by different users. Researchers have undertaken manual analyses of transaction logs and considered changes in both time stamps and the meaning of user queries to demarcate individual searches. Hildreth (1985) and Kurth (1994) commented on these and other disadvantages of transaction logging. To overcome disadvantages of this data collection method, researchers in the United Kingdom advocated the utilization of a frontend system with four capabilities: (1) full-screen logging, (2) playback facility, (3) pre- or post-search, online or offline questionnaire administration, or (4) in-search questionnaire administration (Hancock et al. 1990).

This chapter summarizes what has been learned about the subject terms users enter into online catalogs. Findings are given in the form of generalizations about the subject terms users enter into online catalogs. Studies were limited to those in which researchers selected a random sample of user queries from transaction logs or selected user queries from a transaction log recorded over a certain time period. Studies in which researchers gave respondents a search task to perform and recorded their activity on a transaction log were not included because of the obtrusiveness of this data collection approach.

2.1.1 What Have We Learned about User Queries for Subjects?

In *Using subject headings for online retrieval* (1994), Karen Drabenstott and Diane Vizine-Goetz summarized what had been learned to date about user queries for subjects from online catalog use studies in the form of twenty generalizations. This section gives brief discussions of generalizations. It cites studies that support or fail to support generalizations in the form of study numbers and page numbers; the former correspond to the studies listed in the "References" section that concludes this chapter. (The reference lists of studies that support or fail to support generalizations include a few more studies that have been published since the completion of *Using subject headings for online retrieval*.)

Generalization	Studies
1. A small proportion of the terms users enter into online catalogs are not legitimate subject queries.	Supporting: 16:66; 8:48; 23:76; 3:51; 15:172; 19:270; 12:400; 24:53; 1:68; 5:160–3.

Queries that were not legitimate subject queries were random configurations (///,HJNVM), data entry errors, and expletives. Such activity could be exploratory, accidental, or indicative of the frustration users are experiencing with their online search. Researchers have used different terminology to describe these terms, e.g., malicious entries, garbage, entry errors, graffiti, identifiable obscenities, questions, and comments. The important point about this activity was that it occurred with much less frequency than the entry of legitimate subject queries. For example, such activity was only 0.4% of the random sample of subject queries extracted from transaction logs at Northwestern University (Lester 1989, 172). Thus, online catalog searchers understood enough about the system to enter a legitimate subject search term, i.e., a term that had the potential to retrieve bibliographic records.

Generalization	Studies
2. Users include punctuation in their queries.	Supporting: 8:50; 23:78; 15:184.

Examples of such punctuation were possessive forms with an apostrophe, acronyms with periods in between letters, hyphenated words and phrases, and inverted phrases entered with an intervening comma.

Generalization	Studies
3. Users enter phrases without intervening spaces.	Supporting: 8:48; 13:6; 23:76; 1:68; 5:176.
4. Users misspell words.	Supporting: 16:66; 8:48; 13:4; 23:76; 15:194, 197; 19:170; 12:400; 24:53; 11:28; 2:41; 1:68; 5:175–6; Failing to support: 3: 44.

Spelling errors — substituting, inserting, transposing, and omitting one or more letters — prevented user queries from matching words and phrases indexed in online catalogs. Although spelling errors occur, they were not very prevalent in user queries. For example, about 5% of the subject queries examined in separate studies by Lester (1989, 194, 197) and Jones (1986, 4) contained spelling errors.

Generalization	Studies
5. Users enter abbreviations.	Supporting: 8:48; 23:76; 3:44.

Abbreviated terms were not that common; in fact, Carlyle (1989, 44) only found one such term in her sample of 161 subject terms.

Generalization	Studies
6. Users preface their queries with an initial article.	Supporting: 12:400; 2:41; 1:68.

Initial articles can prevent systems from retrieving bibliographic records unless systems are programmed to delete initial articles from user queries.

Generalization	Studies
7. Users enter queries in singular form when the authorized form in LCSH is plural and vice versa.	Supporting: 3:44; 2:44; 5:177–80.

Barrett and Maticka (1990, 44) gave examples of singular subject headings ("Kimono" and "Wheel") for which users entered plural forms and of plural subject headings ("Apples" and "Kidney beans") for which users entered singular forms.

BEST COPY AVAILABLE

Generalization	Studies
8. Users enter queries containing a date that differs in form from the date in the subject heading in the LCSH system.	Supporting: 3:44.

An example was the user query containing the date "20th century." (LCSH represents this date as "Twentieth century.")

Generalization	Studies
9. Users enter queries containing a different suffix than the subject heading in the LCSH system.	Supporting: 3:44.

An example was the user query containing the suffix "computer programming" that was different from the suffix of the subject heading based on LCSH, viz. "Computer programs."

Generalization	Studies
10. Users enter queries using the online catalog's subject search capability that are probably better suited to the catalog's author or title capability.	Supporting: 19:270; 12:400; 24:53; 5:159–61.

Users who entered such queries could have been unsure which searching capability they should use to enter their query (so they used all of them) or they could havebeen surveying results to determine which capability provided a manageable number of useful retrievals. Fortunately, such queries were not very common.

Generalization	Studies
11. Users enter queries subject queries that match *see* references in the Library of Congress Subject Headings.	Supporting: 16:66; 15:239; 2:41; 5:169.

Lester (1989, 239) reported that 9.8% of subject queries entered into the NOTIS online catalog at Northwestern University matched *see* references in LCSH. Corrections to spelling errors resulted in an additional 0.8% of matches.

Generalization	Studies
12. Users enter queries subject queries that exactly match authorized headings in the Library of Congress Subject Headings (LCSH).	Supporting: 15:172; 6:108; 2:41; 5:167–70, 187–8.

Lester (1989, 172) reported that 40% of the subject queries entered by searchers of the NOTIS online catalog at Northwestern University were exact matches of the subject headings printed in *LCSH*. (She disregarded capitalization and punctuation, i.e., commas, periods, dashes, hyphens, apostrophes, semicolons, and parentheses, in the matching process.)

Generalization	Studies
13. Users enter subject queries that exactly match subject headings in a bibliographic database based on the LCSH system.	Supporting: 16:66; 3:44; 19:269; 12:399; 24:53; 2:41.

Percentages of matching queries ranged from a low of 18% (Markey 1984, 66) to high of 48% (Peters 1989, 269). The different methods used in the five studies to arrive at percentages of matched queries could explain the wide disparity in percentages.

Generalization	Studies
14. Users that are exact matches of the controlled vocabulary are likely to be shorter than nonmatches.	Supporting: 16:67, 70. Failing to support: 15:176.

Lester (1989, 176) reported that exact matches in her study were the same length (1.7 words) as nonmatches. The nonmatches in Markey's (1984, 67, 70) study were longer (2.3 words) than exact matches (1.5 words).

Generalization	Studies
15. The majority of the subject queries users enter into online catalog's using the system's subject searching capabilities are for topical subjects.	Supporting: 15:186; 5:158–9.

Over 70% of the queries in Lester's (1989, 186) study were for topical subjects. A surprising high percentage (44%) of these queries matched subject headings printed in *LCSH*. The percentage (62.2%) reported by Drabenstott and Vizine-Goetz was comparable to the percentage reported by Lester.

Generalization	Studies
16. Queries for names or combinations of names and other words are less likely to match subject headings printed in *LCSH* than queries for topical subjects.	Supporting: 15:188; 5:225–9.

Of the nearly 29% of user queries in Lester's (1989, 188) study that contained personal, corporate, or geographic names, only thirty percent were successful in matching a subject heading printed in *LCSH*. The high percentage of nonmatches could be explained by the exclusion from LCSH of most headings for personal and corporate names and geographic places that name jurisdictional units. (Such headings are more likely to found in the Library of Congress Name Authority File (LCNAF) than in LCSH.)

Generalization	Studies
17. User queries are sometimes an amalgamation of two or more subject headings from the LCSH system.	Supporting: 3:44.

An example was the user query "crystallography geometry" that matched two separate subject headings printed in *LCSH*: "Crystallography" and "Geometry."

Generalization	Studies
18. The retrieval processes in online catalogs can increase the match success rate between user-entered terms and subject headings based on the LCSH system.	Supporting: 15; 5:165–240.

Lester (1989) tested this assertion using twenty-two different retrieval processes on user queries to effect matches of subject headings. Of the many processes, she demonstrated that right truncation, string searching with adjacency, and keyword searching with an implicit Boolean "and" operator significantly improved match success.

Based on the results of an empirical study of user queries, Drabenstott and Vizine-Goetz (1994) concluded that the subject terms users entered into online systems possessed certain characteristics that revealed the subject searching approaches most likely to produce useful titles on the topics users sought. They suggested that computers systems could be programmed to identify many of these characteristics without the aid of human intermediaries, and, thus, respond to user queries with subject searching approaches likely to retrieve useful titles.

Generalization	Studies
19. Users enter queries containing language not used in subject headings in the LCSH system.	Supporting: 3:44; 2:41; 5:214–5.

Between two and five percent of user queries contained language not used in the LCSH system.

Generalization	Studies

20. The subject queries users enter into online systems can be used as see references to authorized subject headings.	Supporting: 22:68.

Designers of the experimental Okapi online catalog in the United Kingdom programmed Okapi to find certain words and phrases in user queries and search for them and their equivalent terms.

2.1.2 What Have We Learned about the Names Users Enter?

In *Using subject headings for online retrieval* (1994), Karen Drabenstott and Diane Vizine-Goetz summarized what had been learned to date about the names users enter into online catalogs in the form of five generalizations. This section gives brief discussions of generalizations. It cites studies that support or fail to support generalizations in the form of study numbers and page numbers; the former correspond to the studies listed in the "References" section that concludes this chapter. (The reference lists of studies that support or fail to support generalizations include a few more studies that have been published since the completion of *Using subject headings for online retrieval*.)

Generalizations were drawn from studies of name access points entered by users conducting searches for known-items. Findings about the personal, corporate, and geographic names users enter in known-item searches can be generalized to subject searches for names because many systems required users to enter name access points for known-item searches in the same form as name access points for subject searches in order to retrieve bibliographic records. Generalizations #2 about punctuation and #4 about misspelled words from the subjects section (section 2.1.1) were also true for name access points. Generalization #10 from the subjects section was true with a minor change, that is, users entered queries using the online catalog's personal name as author search capability that were probably better suited to other catalog capabilities.

Generalization	Studies
1. Users enter personal names in direct form when systems require the entry of inverted forms and vice versa.	Supporting: 4:32; 20:8; 8:51; 19:270; 12:400; 24:52; 11:27; 1:65, 67; 5:224.

Several studies drew this conclusion about the personal-name access points users entered into online catalogs. In Dickson's study (1984, 32), the percentage of personal names with given names entered first was as high as 37%.

Generalization	Studies
2. Users enter two personal names at the same time.	Supporting: 4:32; 20:8; 24:52; 11:27; 1:65; 5:225.

Users occasionally searched for material by or about more than one person. Examples of such queries were "berger, peter and luckman, thomas," "leopold and loeb," and "nietzche and kierkegard."

Generalization	Studies
3. Users enter personal names with other words.	Supporting: 20:8; 15:186; 5:227–9.

Drabenstott and Vizine-Goetz (1994, 159, 227) reported that less than 1% of user queries contained elements for personal names and other words and gave examples: "paintings of pollack" [sic], "descartes future prediction," and "clarence darrow's relegious [sic] views."

Generalization	Studies
4. The middle names or initials users include in their queries for personal names are sometimes counterproductive in helping them find the heading used in the catalog.	Supporting: 4:32; 23:80; 19:270; 12:400.

Of several reasons why the inclusion of middle names or initials was counterproductive, most centered on characteristics of the system into which the query was entered. For example, if the indexed heading had no initial and the user-entered term had an initial, the system placed the user in an alphabetical index of personal names following the name desired by the user and offered no capability for browsing backward in the alphabet.

Generalization	Studies
5. Retrieval processes in online catalogs can increase the match success rate between user-entered names and the access points indexed in bibliographic databases much more than the LCNAF.	Supporting: 20:15; 15:258; 5:230–3.

Taylor (1984, 15) and Lester (1989, 258) concluded that name authority records would help users find the authorized headings for names for only 6% and 1% of the user-entered terms in their respective studies. According to Taylor, a computer program to invert the elements in user-entered names that produced no retrievals would have helped users find an appropriate name heading for 22% of their access points. Both researchers recommended a right truncation program and Lester recommended string and keyword searching approaches for names. Drabenstott and Vizine-Goetz (1994, 230–3) took a different approach; they advocated that systems prompt users for elements of personal-name queries and submit queries to subject searching approaches depending upon the ability of certain approaches to produce retrievals for query elements.

2.2 Subject Searching Improvements

The objective of researchers who conducted studies of the subject terms users entered into online catalogs was to recommend improvements to the subject searching capabilities of online catalogs. Their calls for subject searching improvements were joined by other experts and spokespersons who were thinking about, commenting on, and analyzing research findings. In preparation for her dissertation research, Lester (1989, 58–98) reviewed recommendations for subject searching improvements made by both online catalog researchers and experts such as the following:

- Automatic detection and correction of spelling errors.
- Automatic detection and correction of search format errors.
- Automatic replacement of user queries with *see* references from LCSH.
- System-supplied truncation.
- Boolean-based, keyword searching of subject-rich fields of bibliographic records.

At the conclusion of her empirical study of the subject terms users enter into online catalogs, Lester (1989, 267) recommended "the retrieval processes of right truncation, string searching, and keyword searching" because "each makes significant improvements in match success with the Library of Congress subject headings" but she did not specify which retrieval process was better suited to which user queries.

Such lists of recommendations are limited for several reasons. They do not tell us which improvement will improve the prospects for success for the greatest number of queries entered into online catalogs. Lists do not tell us whether there are certain characteristics of user queries that make them better suited for particular subject searching improvements. They also do not tell us which subject searching capabilities in our existing online catalogs are currently working satisfactorily with user queries and the characteristics of these queries.

2.3 Subject Searching Improvements to Provide Useful Information

While previous studies increase our knowledge of user queries and provide several recommendations for improving subject searching in online catalogs, they leave a key question unanswered, viz. how can online systems respond to user queries with the subject searching approach most likely to succeed in providing relevant information?

The answer to this question was the objective of an empirical study of the subject terms users enter into online catalogs. The findings of this study were published in the research monograph *Using subject headings for online retrieval* (Drabenstott and Vizine-Goetz, 1994). Since the empirical study's findings were the impetus for the research described in this report, we felt it was important to highlist findings of the empirical study in this report. Findings are summarized in subsections of section 2.3.

2.3.1 Computer and Manual Analyses of Subject Queries

In the empirical study, the researchers extracted 54,429 subject queries entered by end users from the transaction logs of three online catalogs: (1) LS/2000 at the University of Kentucky, (2) ORION at UCLA, and (3) SULIRS at Syracuse University. They submitted extracted terms to a computer analysis that involved three successive comparisons to determine how closely user queries for subjects matched the established headings and *see* references in the machine-readable Library of Congress Subject Headings (LCSH-mr).

Almost 25% of subject queries were exact matches of established headings and *see* references in LCSH-mr. (Exact matches disregarded capitalization and punctuation.) Only one percent of subject queries matched normalized forms of established headings or *see* references in LCSH-mr. (Normalization disregarded capitalization, punctuation, stopwords, and word order in queries, and LCSH headings and *see* references.) About 14% of user queries were keyword matches of words in established heading and *see* reference fields of LCSH-mr records. (Keyword matches required every word in the subject query to match keywords in established heading and *see* reference fields of LCSH-mr records). When the three comparisons were performed and matches discarded, about 40% of user queries remained.

The computer analysis gave the researchers no opportunity to determine answers to qualitative questions such as the characteristics of remaining queries and the relevance of matching headings and *see* references. Thus, they conducted a manual analysis of subject terms to answer the following qualitative questions:

1. How closely does the query match the catalog's controlled vocabulary?
2. If the matching term is satisfactory for expressing the user's topic of interest, what system capability would quickly and efficiently deliver the user to the satisfactory term?
3. If the matching term is not satisfactory, what controlled vocabulary term is satisfactory for expressing the user's topic of interest and what system capability would quickly and efficiently deliver the user to the satisfactory term?

The researchers divided queries into groups for topical subjects, geographical names, corporate names, personal names, and various combinations of topical subjects and different types of names. They used a category scheme devised by Carlyle (1989) to determine whether the initial user queries in 1,503 subject searches were exact, partial, keyword, or nonmatches of the established headings and *see* references in LCSH-mr.

2.3.2 Key Findings of the Empirical Study

Five key findings of the empirical study of user queries are enumerated below along with system capabilities that were supported by the particular finding.

1. Topical subject queries and geographical name queries register high percentages of exact matches.

The matching term often was posted in hundreds of bibliographic records. For example, user queries such as "art," "computers," "photographs," "united states," and "great britain" were exact matches of LCSH-mr headings that retrieved hundreds of bibliographic records in many library databases. The number of retrieved bibliographic records would increase into the thousands if subdivided forms of these headings were retrieved.

To perform retrieval quickly and efficiently, online catalogs need a variation of the alphabetical approach to subject searching. Such a variation would anticipate the user's selection of the exact match from an alphabetical list, and thus begin with a report of the results of such exact matches, summarize the retrieval of subdivided forms, and give users the option to browse related terms from LCSH-mr if available.

2. Partial matches are: (1) most combinations of queries for topical subjects and names, and (2) between one-quarter and one-third of queries for topical subjects and geographical names, respectively.

Partial matches were satisfactory representations of the topics of interest users had in mind only when queries matched the initial words in a longer assigned heading or LCSH-mr *see* reference. The alphabetical approach would be helpful for placing users in the neighborhood of potentially relevant controlled vocabulary terms. Different subject searching approaches would be required for other partial matches.

3. Most user queries would be satisfied by the alphabetical or keyword approaches implemented in existing online catalogs.

Existing online catalogs feature the following subject searching approaches: (1) alphabetical, (2) keyword-in-main-heading, (3) keyword-in-subdivided heading, and (4) keyword-in-record. Except for a few queries, most queries would retrieve useful material using one or several of these approaches. Although many online catalogs feature more than one of these approaches, they do not provide the user with guidance as to which approach will provide useful results in response to a particular query.

4. The subject terms users enter into online systems possess certain characteristics that reveal the subject searching approaches most likely to succeed at providing useful information on the topics users seek.

Examples of these characteristics were the number of words in user queries, the extent to which user queries matched controlled vocabulary terms, and their ability to produce retrievals in response to certain subject searching approaches. Online systems could help users choose among existing subject searching approaches by determining how closely their subject queries match assigned subject headings in library databases and produce retrievals.

5. Users enter the various elements of personal name queries in an unpredictable sequence.

The order in which users entered the elements of personal name queries was unpredictable even when users entered personal name queries in systems that required a particular form of name, e.g., direct form or inverted form. Systems need a separate subject search option for personal names. When users select such an option, systems should prompt users to enter the four elements of personal name queries, i.e., last name, first name, middle name or initial, and topic, or as many of the elements as they know. The particular elements users provide and the system's ability to find these elements in personal name headings determines the subject searching approaches used by the system.

2.4 Subject Searching Functionality and Search Trees

2.4.1 Subject Searching in Operational Online Catalogs

Operational online catalogs feature keyword and alphabetical approaches to the subject queries users enter. The former approach gives a report of the number of assigned subject headings or cataloging records bearing words in the user query. The latter approach produces a list of assigned subject headings in the alphabetical neighborhood of the user query.

Consider the user query "aids." When "aids" is entered through keyword approaches, systems retrieve thousands of cataloging records but most retrieved records are not relevant because retrievals are made for the word "aids" in the subdivision "Audio-visual aids" that is connected to hundreds of subject headings on many different topics. Results of the alphabetical approach would be much more useful because the list of subject headings in alphabetical neighborhood of "aids" would probably include the relevant heading "AIDS (Disease)" and several subdivided forms of this heading, e.g., "AIDS (Disease) — Alternative treatment" and "AIDS (Disease) — Biography." Although both approaches retrieve relevant cataloging records, most records retrieved through the alphabetical approach are going to be specifically about the AIDS disease because they are retrieved through the relevant subject heading for this topic.

How would online catalog users know whether to use keyword or alphabetical approaches for this query? Do users know the differences between approaches? Although the majority of online catalogs in American libraries offer more than one approach to subject searching, they give users little, if any, guidance as to which approach is better suited to their queries. Some catalogs feature as many as five subject searching approaches! Users could enter their queries using several approaches and evaluate the results for themselves. But how many users would be patient enough to review the results of more than one subject searching approach to determine the most appropriate system response?

2.4.2 Search Trees: An Innovative Approach to System Design

Search trees hold much promise for assuming the burden of determining which subject searching approach is likely to produce useful information for user queries. The designers of the Okapi experimental online catalog first defined search trees as "a set of paths with branches or choices, which enables the system to carry out the most sensible search function at each stage of the search" (Mitev, Venner, and Walker 1985, 94). The search trees they implemented in OKAPI "evolved through a process of discussion and trial and error" and placed more emphasis on searching the titles than the subject headings in OKAPI's cataloging records because only half of these records contained subject headings (Mitev, Venner, and Walker 1985, 94).

Some online catalogs have subject searching routines that resemble search trees. For example, the online catalog of the University of Illinois at Urbana-Champaign responds to user queries for subjects with keyword searches of assigned subject headings. When users terminate searches, the system prompts them to continue and gives the results of a title-keyword search (Hildreth 1989, 86–7). The Illinois online catalog always performs keyword searches of subject heading fields before title-keyword searches because the former consumes fewer system resources than the latter.

The search trees tested in this research effort were the result of the empirical study of the subject terms users entered into online catalogs (Drabenstott and Vizine-Goetz 1994). These search trees emphasized subject headings because the vast majority of cataloging records created by American libraries are assigned subject headings based on the Library of Congress Subject Headings (LCSH) (O'Neill and Aluri 1979, 5).

The search trees exemplified the searching strategies used by expert search intermediaries. Intermediaries use controlled vocabulary because it yields relevant output. When controlled vocabulary is not available to express user queries, they conduct free text searches of titles and abstracts to retrieve a few relevant records, review results to find relevant controlled vocabulary, and then incorporate such vocabulary into the ongoing search. The search trees performed in a similar manner. They invoked searching approaches that looked for matches of user queries in subject heading fields of cataloging records before enlisting keyword search approaches that looked for matches in title fields or in a combination of title and subject heading fields. Chapter 4 gives a full description of the search trees tested in this study.

2.5 Chapter Summary

This chapter presented generalizations about user queries for subjects. Generalizations were derived from studies of user queries in which researchers sampled user queries from transaction logs or selected user queries from a transaction log recorded over a certain time period.

Although previous studies increase our knowledge of user queries and provide several recommendations for improving subject searching in online catalogs, they leave a key question unanswered, viz. how can online systems respond to user queries with the subject searching approach most likely to succeed in providing relevant information?

The answer to this question was the objective of an empirical study of the subject terms users enter into online catalogs (Drabenstott and Vizine-Goetz 1994).

The findings of the empirical study were the foundation for the new design of subject access to online catalogs that included search trees. Search trees are a set of paths with branches or choices that enable systems carry out the most sensible search function at each stage of the search. The search trees tested in this study exemplified the searching strategies used by expert search intermediaries who favor searches of controlled vocabulary fields over other content-rich fields of cataloging records. Search trees hold much promise for assuming the burden of determining which subject searching approach is likely to produce useful information for user queries. This report describes a test of the new subject access design in the ASTUTE experimental online catalog.

References

1. Ballard, Terry, and Jim Smith. 1992. "The human interface: an ongoing study of OPAC usage at Adelphi University." In *Advances in online public access catalogs*, edited by Marsha Ra, 58–73. Westport, CT: Meckler.
2. Barrett, Beverly, and Margaret Maticka. 1990. "An analysis of user failure in subject searching an online catalogue." In *Garbage in, garbage out: The need for quality in the age of automation; Australian Library and Information Association 8th national cataloguing conference*, 38–49. Adelaide: Auslib Press.
3. Carlyle, Allyson. 1989. "Matching LCSH and user vocabulary in the library catalog." *Cataloging & Classification Quarterly* 10, 1/2: 37–63.
4. Dickson, Jean. 1984. "An analysis of user errors in searching an online catalog." *Cataloging & Classification Quarterly* 4, 3 (Spring): 19–38.
5. Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential*. San Diego: Academic Press.
6. Freiburger, Gary, and Marjorie Simon. 1983. "Patron use of an online catalog." In *Crossroads: Proceedings of the first national conference of the Library and Information Technology Association*, edited by Michael Gorman, 106–11. Chicago: American Library Association.
7. Hancock, Micheline, et al. 1990. *Evaluation of online catalogues: an assessment of methods*. London: British Library. British Library Research Paper no. 78.
8. Henty, Margaret. 1986. "The users at the online catalogue: a record of unsuccessful keyword searches." *LASIE* 17, 2 (September/October): 4–52.
9. Hildreth, Charles R. 1989. *Intelligent interfaces and retrieval methods for subject searching in bibliographic retrieval systems*. Washington, DC: Library of Congress. Advances in Library Information Technology 2.
10. Hildreth, Charles R. 1985. "Monitoring and analyzing online catalog user activity." *LS/2000 Communiqué*, pp. 3–6.
11. Holmes, David, and Derrick Bulger. 1988. "A day in the life of a public terminal — a transaction analysis of an online catalogue terminal in a bilingual environment." *Canadian Journal for Information Science*, 13, 3/4: 21–33.

12. Hunter, Rhonda N. 1991. "Successes and failures of patrons searching the online catalog at a large academic library: a transaction log analysis." *RQ* 30, 3 (Spring): 395–402.

13. Jones, Richard. 1986. "Improving Okapi: transaction log analysis of failed searches in an online catalogue." *Vine* no. 62: 3–13.

14. Kurth, Martin. 1994. "The limits and limitations of transaction log analysis." *Library Hi Tech* 11, 2: 98–104.

15. Lester, Marilyn Ann. 1989. *Coincidence of user vocabulary and Library of Congress Subject Headings: experiments to improve subject access in academic library online catalogs.* Ph.D. dissertation, University of Illinois at Urbana-Champaign.

16. Markey, Karen. 1984. *Subject searching in library catalogs: before and after the introduction of online catalogs.* Dublin, Ohio: OCLC.

17. Mitev, Nathalie, Gillian Venner, and Stephen Walker. 1985. *Designing an online public access catalog.* London: British Library. Library and Information Research Report 39.

18. O'Neill, Edward T., and Rao Aluri. 1979. *Subject heading patterns in OCLC monographic records.* Columbus, Ohio: OCLC, Inc. OCLC Research Report Series OCLC/OPR/RR–79/1.

19. Peters, Thomas A. 1989. "When smart people fail: An analysis of the transaction log of an online public access catalog." *Journal of Academic Librarianship* 15, 5 (November): 267–73.

20. Taylor, Arlene G. 1984. "Authority files in online catalogs: An investigation of their value." *Cataloging & Classification Quarterly* 4, 3 (Spring): 1–19.

21. Van Pulis, Noelle, and Lorene E. Ludy. 1988. "Subject searching in an online catalog with authority control." *College & Research Libraries* 49 (6): 523–33.

22. Walker, Stephen, and Richard M. Jones. 1987. *Improving subject retrieval in online catalogues: 1. Stemming, automatic spelling correction and cross-reference tables.* London: British Library. British Library Research Paper no. 24.

23. Walter, Dennis R. 1987. "The user at the online catalogue: A record of unsuccessful keyword searches — another case study." *LASIE* 18, 3 (November/December): 74–81.

24. Zink, Steven D. 1991. "Monitoring user search success through transaction log analysis: the WolfPAC example." *Reference Services Review* 19, 1 (Spring): 49–56.

3 Machine-readable Data for System Development

3.1 Introduction

Chapters 3–5 highlight the development of the ASTUTE experimental online catalog. This chapter begins with a description of the computer equipment used for ASTUTE development. It features a description of the machine-readable bibliographic and authority data that serve as the foundation for ASTUTE's searchable databases and dictionaries.

3.2 Computer Equipment

ASTUTE (A Search Tree Underlying the Experiment) was programmed on a stand-alone Gateway 2000 486, 33 MHz, IBM-compatible microcomputer, with 8 megabytes of RAM and a VGA color monitor. The operating system was MS-DOS version 5.0. A dot-matrix printer and a mouse were attached to the microcomputer for use by ASTUTE project staff during development work and end users during online retrieval tests.

The ASTUTE project team used this computer equipment and a connection to the university' campus network for various tasks, e.g., transmitting data on magnetic tape to the microcomputer's hard disk, transmitting transaction log files of end user search activity between participating institutions. When ASTUTE was installed in the libraries of University of Michigan-Dearborn (UM-D) and Earlham College, the system did not require a network connection. It resided entirely on the Gateway microcomputer. Participating library staff monitored the system, performed daily system backups, and periodically used the microcomputer equipment and network connections in their institutions to transmit transaction log files of end user search activity to the ASTUTE project team in Ann Arbor.

3.3 Machine-readable Bibliographic Data

The databases of the ASTUTE experimental online catalog were created from two data sources: (1) machine-readable cataloging (MARC) records for bibliographic data from the two participating libraries in selected subject areas of the Library of Congress Classification (LCC), and (2) MARC records for subject authority data from the compact disk-based product CD/MARC Subjects distributed by the Library of Congress. The number and subject areas of MARC bibliographic records were:

1. Mardigian Library of the University of Michigan-Dearborn: 14,686 bibliographic records in Computer Science (QA76's) and Technology (T–TX).
2. Lilly Library of Earlham College: 11,976 bibliographic records in American History (E1–F1199).

The ASTUTE project team did not combine bibliographic records into a single database. Rather, the team used the two libraries' bibliographic records to create separate, searchable databases on computer science and technology for UM-D and on American history for Earlham College, respectively. During online retrieval tests at Mardigian Library and Lilly Library, users searched a database of materials available at the library at which they were searching ASTUTE.

3.3.1 Obtaining Bibliographic Data from Participating Libraries

UM-D provided bibliographic records in USMARC format on two magnetic tapes in ASCII format. To transfer records from tape to the Gateway microcomputer's hard disk, the tape was mounted at the University of Michigan (UM) computer center in Ann Arbor, copied to a file on the mainframe, and the file was downloaded to the Gateway microcomputer's hard disk. All UM-D bibliographic records were accepted for processing into ASTUTE.

An ASTUTE project team member traveled to Earlham College to download USMARC format bibliographic records from Lilly Library's Marcive Public Access Catalog (PAC) to floppy disk. The team member used Marcive PAC's call number browsing capability to produce brief title displays in call number order, highlight titles in blocks of two hundred, and, for highlighted titles, download USMARC-format bibliographic records to floppy disk in ASCII format. Downloaded records on floppy disk were loaded directly into the Gateway microcomputer in the project team office in Ann Arbor. Downloaded records included duplicate records because of the task-intensive manual downloading process; consequently, the project team included a routine in the conversion program described in section 3.3.2 to avoid writing duplicate records. Except for duplicates, all Earlham records were accepted for processing into ASTUTE. Since bibliographic records from UM-D and Earlham were USMARC-format records, the ASTUTE project team used similar computer programs to process bibliographic data into ASTUTE for database building.

3.3.2 Converting Bibliographic Data

After ASCII-format bibliographic records were loaded into the Gateway microcomputer's hard disk, the ASTUTE Project Team wrote a bibliographic conversion program to convert ASCII-format bibliographic records contributed by UM-D and Earlham into databases (or files) for local processing format. The conversion programs were written using low-level file access functions in FoxPro. The programs created seven FoxPro database tables and two temporary databases for listing subject headings (USMARC tags 650 and 651). Additional database tables were created from the original seven databases with several programs, depending on the search criteria of the individual search trees.

Programs written for converting UM-D and Earlham records were very similar. The only differences were connected with parameters of Library of Congress call number fields and the ASCII characters beginning each bibliographic record in the file.

USMARC-format records for bibliographic data consist of a 24-character leader (or header), a directory record bearing a series of twelve characters for each tag within the record that provide the map for finding data fields, and the variable-length data fields in the record. Conversion programs first processed the directory records for each bibliographic record. They skipped the leader in each record (i.e. the first twenty-four characters) and positioned the file pointer at the beginning of directory records. Each directory record was read in individually twelve characters at a time into a memory variable. Programs checked the first character of the memory variable to see if the character was ASCII character 30 which indicated the end of the directory. If the end of the directory was not found, programs divided directory records into the USMARC tag number, the length of the record, and the starting position of the record and these were stored in a directory array or table in memory. The program continued reading and storing directory records into an array until ASCII character 30 was found indicating the end of the directory. When ASCII character 30 was found, the file pointer was set back eleven characters to the beginning of the variable length data fields. The program then sorted the table of directory records on the order of the starting position of the variable-length fields in the file.

Conversion programs processed variable-length fields in bibliographic records after the directory records were stored in the directory array. They checked the USMARC tag in the array to see if the tag was to be converted into the FoxPro database. If the field was needed, the file pointer was placed at the starting position in the file of the variable-length field, and the total number of characters for the record were read into a memory variable. Conversion programs searched for ASCII character 31 which indicated subfield divisions within a field. Conversion programs assigned each bibliographic record a unique number to index on a unique field for data lookups. Programs distributed bibliographic record data in desired fields and subfields into five FoxPro bibliographic databases and the two temporary databases for subject headings bearing USMARC tags 650 and 651, respectively. Programs continued processing each element in the directory array until all of the USMARC fields in the directory array were written to the appropriate FoxPro databases.

Conversion programs then continued to read in data one character at a time to find the beginning of the next record. A check was done to see if the end of the file had been found. If this check found the end of the file, processing was finished and if not, the character was checked to see if it was the beginning of the next bibliographic record. ASCII character 29 was the beginning of each UM-D record and ASCII character 10 was the beginning of each Earlham record. When the beginning of the new record was found, the program again skipped the twenty-four character header and continued processing directory records into a new directory array.

The project team designed the bibliographic databases and tested the conversion programs thoroughly to make sure that accurate information was imported into the FoxPro databases on the Gateway microcomputer.

3.3.3 Changing Bibliographic Data

3.3.3.1 Diacritics

ASTUTE was developed using the FoxPro database management system version 2.0 (FoxPro Software, Inc. 1991). FoxPro supports and displays the standard 255 ASCII character set. The standard ASCII characters for many of the diacritics found in the bibliographic records were not displayed as they would be on a system using a different character set. When FoxPro displayed bibliographic data bearing diacritics, it indicated the presence of diacritics by displaying a standard ASCII character in the ASCII 255-character set (i.e., the accent mark sometimes printed above the last "e" in the word resume was displayed as a separate ASCII character). The ASTUTE project team did not delete diacritics from UM-D bibliographic records because diacritics were so uncommon in the UM-D database of computer science and technology records that their presence would not adversely affect retrieval.

Diacritics were much more common in Earlham College bibliographic records on American history. Since the presence of diacritics could adversely affect retrieval, the ASTUTE Project Team wrote, tested, and executed a program that deleted diacritics from bibliographic records.

3.3.3.2 Subject Headings

Mardigian Library (UM-D) staff advised the ASTUTE project team that the subject headings in their library's bibliographic records would contain errors. The project team wrote, tested, and executed a program to produce two alphabetical lists of unique subject headings for topics and geographic names (USMARC tags 650 and 651, respectively). In the process of downloading authority records from CD/MARC Subjects (see section 3.3.4), project team reviewed listed subject headings and annotated the list with corrections. The project team manually corrected erroneous subject headings in UM-D bibliographic records. Frequent errors in main headings and subdivisions were: (1) assignment of *see* references, (2) incorrect tags, (3) incorrect subfield codes, (4) misspellings, (5) abbreviations, (6) canceled subject headings, and (7) canceled subdivisions.

Although Lilly Library (Earlham) staff did not expect many erroneous subject headings, ASTUTE project staff also manually corrected erroneous subject headings in Earlham bibliographic records. Since these records contained many names used as subjects, the project team produced a third alphabetical list for unique names as subjects (USMARC tag 600) and encountered errors they did not encounter with topical subjects, e.g., incorrect dates, open date ranges, and missing qualified names.

3.3.4 Enhancing Bibliographic Data

3.3.4.1 Form Subdivisions

The objective of ASTUTE's exact approach to subject searching (see section 5.3) was to manage lengthy browsing displays under a given search term using broad categories. Broad categories corresponded to subfield codes in subject heading fields, thus, they were limited to three subfield codes: (1) $x for topical subdivisions, (2) $y for period subdivisions, and (3) $z for geographic subdivisions.

The project director and others have called for an editorial review of subdivisions in the *Subject cataloging manual: subject headings* (SCM:SH) for the purpose of creating broad categories (Massicotte 1988; Holly and Killheffer 1982; Drabenstott and

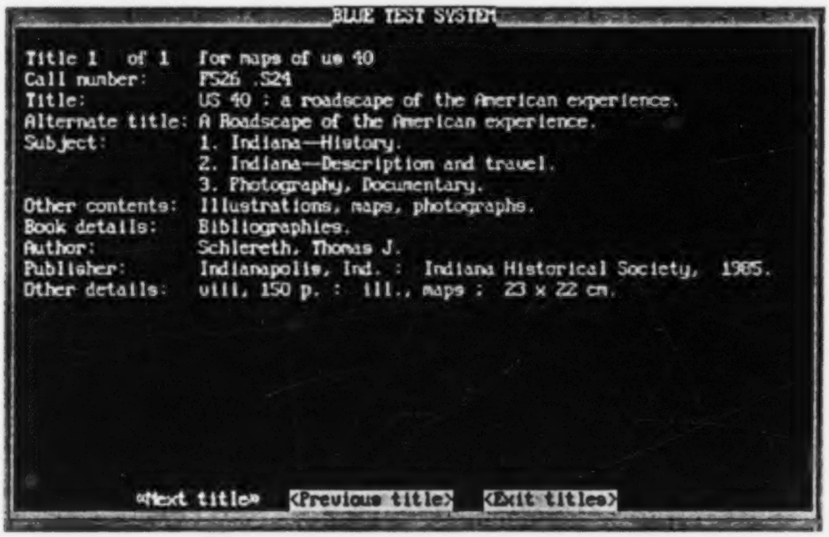
Vizine-Goetz 1990, 13, 1994, 254–60). The editorial review would be a huge undertaking because of the thousands of subdivisions in SCM:SH. In the absence of such a review, the ASTUTE project team created one new broad category for form subdivisions. The team reviewed an alphabetical list of unique topical subdivisions in UM-D and Earlham bibliographic records and reassigned selected subdivisions to a new "local" subfield code ($1). Examples of form subdivisions are "Amateurs' manuals," "Bibliography," "Films," "Maps," and "Study guides." Appendix A gives two separate lists of the form subdivisions in UM-D and Earlham bibliographic records.

3.3.4.2 Control Field Codes

Several coded positions in the 008-control field of USMARC bibliographic records convey information about the physical and bibliographic characteristics of the item. Unfortunately, replacing the codes with control-field codes has been ignored or overlooked in cataloging and in system design (Byrne (1987a, 4).

Realizing the value of coded data in the 008-control field, the ASTUTE project team identified the following four control field elements that would be useful in subject searches and bibliographic record displays: (1) illustration codes (positions 18–21), (2) nature of contents codes (positions 24–27), (3) biography code (position 34), and (4) festschrift code (position 30). The team programmed ASTUTE to convert codes in these control-field positions to the English-language equivalent terms for use in subject searching and displaying bibliographic records. ASTUTE includes English-language equivalent terms from these four elements in keyword-in-record searches for subjects generally and for personal names as subjects. For example, a keyword-in-record search in ASTUTE for the user query "maps of us 40" produced retrievals from several subject-rich fields including the 008-control field. One bibliographic record that was retrieved in the search was *US 40: a roadscape of the American experience.* Figure 3.1 shows this record; it bears the terms "us" and "40" in the title field and "maps" in the "other contents" field.

Appendix B lists English-language equivalent terms and display field labels for the four control field elements ASTUTE used in keyword-in-record searches for subjects generally and for personal names.



```
Title 1   of 1   for maps of us 40
Call number:     F526 .S24
Title:           US 40 : a roadscape of the American experience.
Alternate title: A Roadscape of the American experience.
Subject:         1. Indiana—History.
                 2. Indiana—Description and travel.
                 3. Photography, Documentary.
Other contents:  Illustrations, maps, photographs.
Book details:    Bibliographies.
Author:          Schlereth, Thomas J.
Publisher:       Indianapolis, Ind. :  Indiana Historical Society,  1985.
Other details:   viii, 150 p. :  ill., maps :  23 x 22 cm.
```

Figure 3.1. Title bearing terms for control-field codes

3.4 Subject Authority Data

3.4.1 Obtaining Subject Authority Data

In 1986, the Library of Congress (LC) began distributing subject authority data to subscribing libraries. Libraries responded by requiring subject authority data functionality in system specifications and setting up operations to maintain the currency of subject authority data in their automated systems (Garrison 1991). Vendors of automated library systems responded by enhancing their online catalogs with subject authority data and functionality to handle such data (Drabenstott 1991, 5–6).

Subject authority records have several features to help subject searchers: (1) scope notes to define the use of the subject heading, (2) broader, narrower, and related terms to help users refine their topics of interest, and (3) cross references to guide users from unused words and phrases to authorized subject headings. The ASTUTE project team added subject authority data and functionality into ASTUTE because library staff and catalog users take these features for granted in today's online catalogs.

LC distributes machine-readable subject authority data on magnetic tapes (machine-readable Library of Congress Subject Headings [LCSH-mr]) and on compact disk (CD/MARC Subjects). The ASTUTE project team inquired about the availability of

BEST COPY AVAILABLE

LCSH-mr or CD/MARC Subjects. LC responded by giving the project team a complimentary copy of CD/MARC Subjects.

The ASTUTE project team programmed FoxPro to generate alphabetical lists of unique subject headings for topical subjects and geographic names (MARC tags 650 and 651) in UM-D and Earlham bibliographic records. Team members used these lists to search CD/MARC Subjects. They downloaded subject authority records for subject headings that matched the main subject heading or the main subject heading and one or more subdivisions of listed headings. Table 3.1 cites the number of subject authority records selected and downloaded. The number of matching authority records was smaller than the number of unique subject headings because many subject headings were subdivided, and, thus, had no corresponding records in the subject authority file. Records downloaded from CD/MARC Subjects were USMARC format records for authority data. Records were loaded directly into the Gateway microcomputer in the project team office in Ann Arbor.

Table 3.1. Authority Records Selected

Library	Number of Unique Subject Headings	Number of Matching Authority Records
UM-D	7,792	4,219
Earlham	9,345	3,824

3.4.2 Converting Subject Authority Data

The ASTUTE project team wrote a conversion program to convert the USMARC authority records from ASCII format files to the FoxPro database format. The program contained a routine to avoid writing duplicate records. The conversion program converted the selected subject authority records to five FoxPro databases (or files). The conversion process was very similar to the one used to convert the bibliographic records (section 3.3.2). The ASTUTE project team handled diacritics in the same way as they handled diacritics in bibliographic records (section 3.3.3.1).

3.4.3 Enhancing Subject Authority Data with Narrower Terms

The ASTUTE project team wrote and implemented a program to verify whether broader and related terms in subject authority records were used in participating library bibliographic records. Unused terms were deleted from subject authority records. When broader terms were used in bibliographic records, they were added as narrower terms to the appropriate subject authority record. For example, the subject heading "Archaeology" was a broader term in the subject authority record for "Cliff-dwellings." Since both terms were used in Earlham bibliographic records, the program added "Cliff dwellings" as a narrower term to the "Archaeology" authority record.

3.5 Chapter Summary

Chapter 3 was the first of three chapters highlighting the development of the ASTUTE experimental online catalog. This chapter began with a description of the computer equipment used for ASTUTE development. It featured a description of the machine-readable bibliographic and authority data that served as the foundation for ASTUTE's searchable databases.

The databases of the ASTUTE experimental online catalog were created from two data sources: (1) machine-readable cataloging (MARC) records for bibliographic data from the two participating libraries in selected LCC subject areas, and (2) USMARC records for subject authority data from the compact disk-based product CD/MARC Subjects distributed by the Library of Congress. Details on how the ASTUTE project team obtained bibliographic data (section 3.3) and authority data (section 3.4) were provided.

The team made few changes to bibliographic data: (1) deleting diacritics in Earlham bibliographic records (section 3.3.3.1), and (2) correcting erroneous subject headings (section 3.3.3.2). The few enhancements the team made to bibliographic data were intended to improve certain subject searching approaches: (1) identifying and coding form subdivisions to increase the number of broad categories in exact search from three to four (section 3.3.4.1), and (2) converting control-field codes to English-language equivalent terms to increase the searchable dictionary in keyword-in-record searches (section 3.3.4.2). The team deleted diacritics in subject authority records (section 3.4.2) and enhanced these records with narrower terms (section 3.4.3).

References

Byrne, Deborah. 1987a. "The much-misunderstood MARC fixed field: Part 1 of 3." *Action for Libraries* (February): 4.

Byrne, Deborah. 1987b. "The much-misunderstood MARC fixed field: Part 3 of 3." *Action for Libraries* (April): 3–4.

Drabenstott, Karen Markey. 1991. "Introduction." In *Subject authorities in the online environment*, 1–8. Chicago: American Library Association.

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Drabenstott, Karen Markey, and Diane Vizine-Goetz. 1990. "Search trees for subject searching in online catalogs." *Library Hi Tech* 8, 3: 7–20.

FoxPro Software, Inc. 1991. *FoxPro commands and functions.* Perrysburg, Ohio: FoxPro Software, Inc.

Garrison, William A. 1991. "Practical considerations in using the machine-readable LCSH." In *Subject authorities in the online environment*, 41–55. Chicago: American Library Association.

Holley, Robert P., and Robert E. Killheffer. 1982. "Is there an answer to the subject access crisis?" *Cataloging & Classification Quarterly* 1, 2/3: 125–33.

Massicotte, Mia. 1988. "Improved browsable displays for online subject access." *Information Technology and Libraries* 7, 4 (December): 373–80.

4 Search Trees for Subject Searching

4.1 Introduction

The subject terms users enter into online systems possess certain characteristics that reveal the subject searching approaches most likely to succeed in terms of producing assigned subject headings and bibliographic records on the topics users seek. Computer systems can be programmed to identify many of these characteristics without the aid of a human intermediary. Examples of characteristics are the number of words in user queries, the extent to which user queries match controlled vocabulary terms, and their ability to produce retrievals in response to certain subject searching approaches. Programming computer systems to identify these characteristics would enable systems to respond to the subject queries users enter with the subject searching approaches most likely to succeed at providing relevant information.

The ASTUTE experimental online catalog was composed of two catalogs: (1) Blue Test System in which search trees govern the system's selection of subject searching approaches, and (2) Pinstripe Test System in which subject searching approaches were selected randomly. This chapter describes the search trees of the Blue System that control its selection of subject searching approaches (sections 4.2–4.5). It sets the stage for chapter 5 discussions of subject searching approaches in the Blue and Pinstripe Systems.

4.2 Search Tree Design

Search trees governed the Blue Test System's selection of subject searching approaches. In the Blue System, users did not explicitly choose a particular approach. Rather, the system responded with an approach based on the characteristics of user queries and their ability to produce retrievals. When the system selected an approach yielding zero or too few retrievals, users could continue searching using another approach. Search trees also governed the Blue System's selection of subsequent subject searching approaches.

Six search trees for subject searching are presented in this chapter. These search trees were implemented into the Blue Test System. Search trees incorporated into the Blue Test System emanated from the findings of an empirical study of the subject terms users enter into online catalogs (Drabenstott and Vizine-Goetz, 1994).

4.3 Initial Search Tree

The user's selection of the subject searching option in the Blue Test System set into motion the initial search tree. The initial search tree was a filter because it dispatched user queries to a particular search tree that favored the selection of certain subject searching approaches over others. Figure 4.1 is the initial search tree.

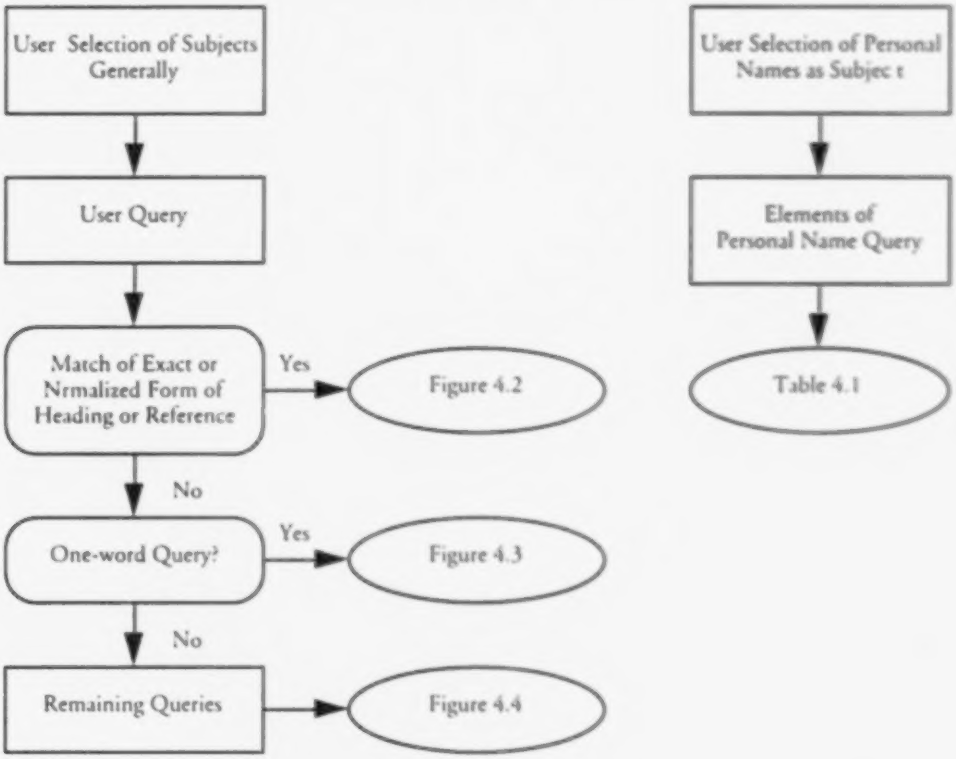

User Selection of Subjects Generally
User Query
Match of Exact or Nrmalized Form of Heading or Reference
Yes
Figure 4.2
No
One-word Query?
Yes
Figure 4.3
No
Remaining Queries
Figure 4.4
User Selection of Personal Names as Subjec t
Elements of Personal Name Query
Table 4.1

Fig. 4.1. Initial search tree

The initial search tree let users distinguish their queries for personal subjects from queries for topical subjects generally and, based on summary characteristics that systems determined about the latter types of user queries, dispatched them to a particular search tree that favored certain subject searching approaches over others. All user queries for subjects generally were candidates for the exact approach. To effect an exact match, the Blue System manipulated user queries in the same way as controlled vocabulary terms were manipulated to establish exact and normalized forms, e.g., ignoring capitalization, removing punctuation and stopwords. In the event an exact match was found, the initial search tree dispatched the query to the search tree that governed the exact approach (figure 4.2).

Unable to find an exact match, the initial search tree dispatched user queries to one of two search trees based on the number of words in queries. One-word queries were given to a search tree that favored the alphabetical approach (figure 4.3). The extent to which remaining queries matched controlled vocabulary terms determined whether they were submitted to a search tree that favored the alphabetical or keyword approaches (figures 4.4 and 4.5, respectively).

The Blue System prompted users entering personal-name queries for the name and topic elements of personal-name queries. Personal-name queries were handled by a search tree that was separate from search trees for subjects generally (Table 4.1). The Blue System chose between alphabetical and keyword approaches depending on the types of elements users entered and the ability of these elements to produce retrievals.

4.4 Search Trees for Subject Queries Generally

The Blue System responded to user queries for subjects generally that matched exact or normalized forms of controlled vocabulary terms with the exact approach. Figures 4.2A and 4.2B depict the search tree that featured the exact approach. Details about the exact approach are given in chapter 5.

The exact approach search tree was split into two parts. Figure 4.2A depicts major events of the exact approach in which the system presented a conceptual map to subdivided forms of the matched heading, and, if available, gave users the option to browse related terms and other information about the matched heading.

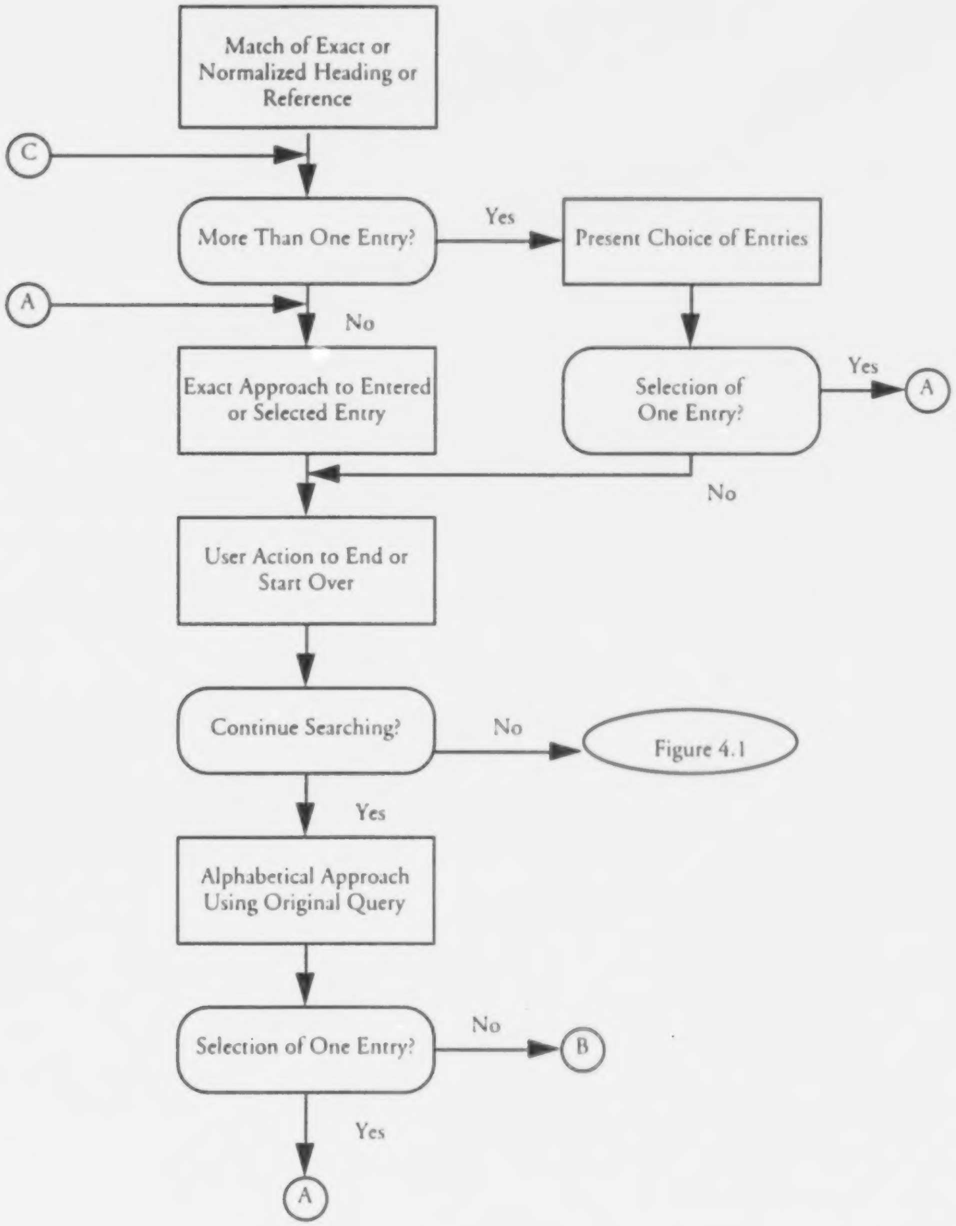

Match of Exact or Normalized Heading or Reference
C
More Than One Entry?
Yes
Present Choice of Entries
A
No
Exact Approach to Entered or Selected Entry
Selection of One Entry?
Yes
A
No
User Action to End or Start Over
Continue Searching?
No
Figure 4.1
Yes
Alphabetical Approach Using Original Query
Selection of One Entry?
No
B
Yes
A

Figure 4.2A. Search tree for the exact approach

Figure 4.2B depicts major events following a user action to start over or end the search. The search tree submitted or the original user query to other approaches beginning with controlled vocabulary approaches. The Blue System continued searching using the alphabetical approach and various keyword approaches beginning with the keyword-in-main-heading search.

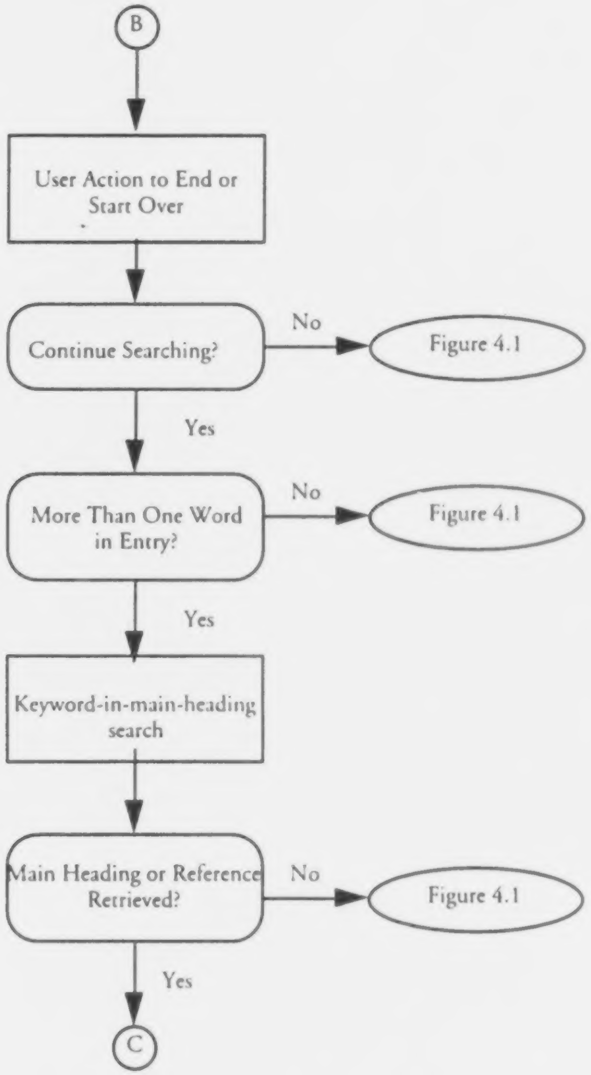

B
User Action to End or Start Over
Continue Searching?
No
Figure 4.1
Yes
More Than One Word in Entry?
No
Figure 4.1
Yes
Keyword-in-main-heading search
Main Heading or Reference Retrieved?
No
Figure 4.1
Yes
C

Figure 4.2B. Search tree for the exact approach (contd.)

Figure 4.2B shows only the keyword-in-main-heading search. This tree could be expanded to include keyword-in-subdivided-heading, title-keyword, keyword in subject heading fields, and keyword-in-record searches.

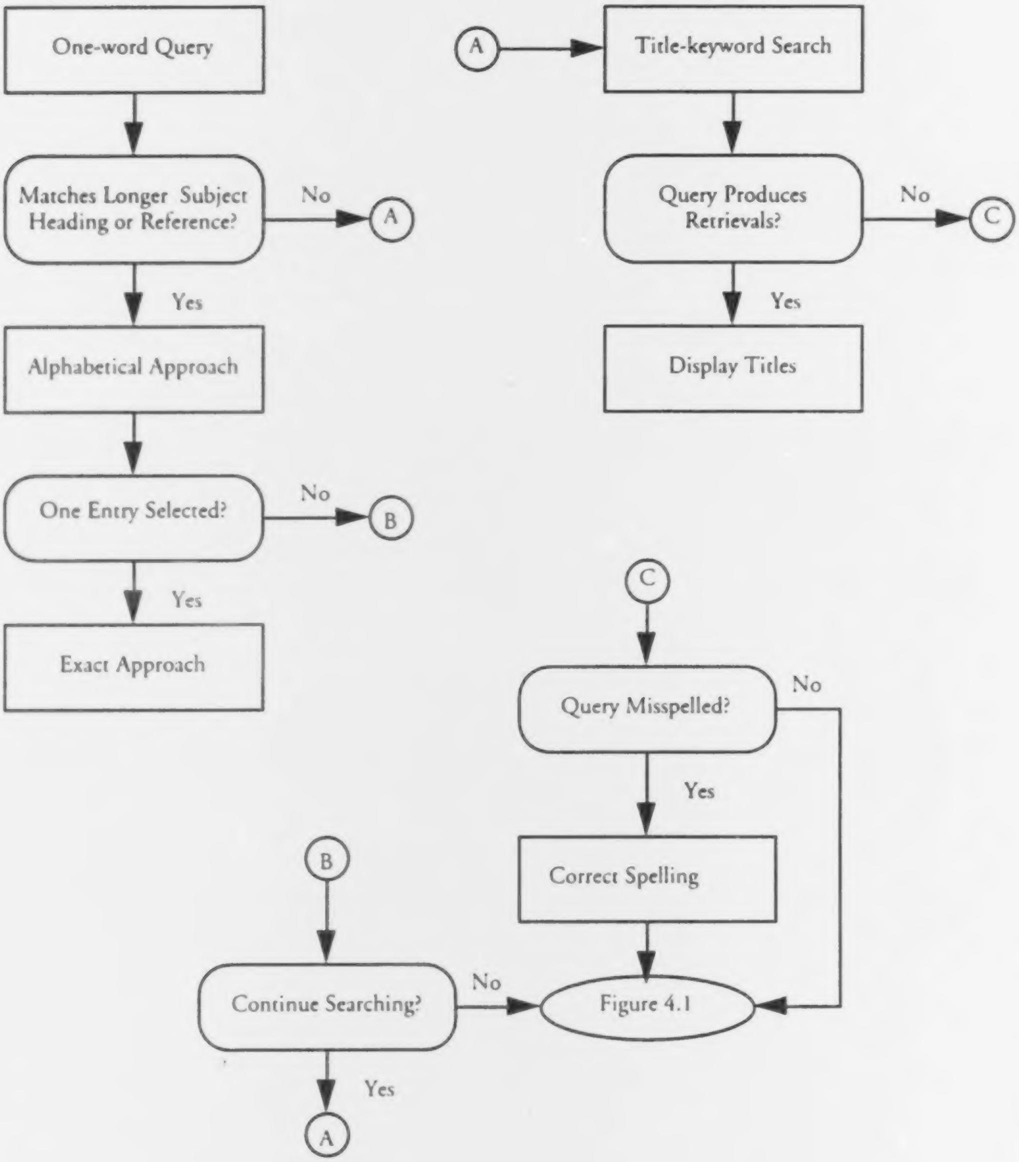

One-word Query
Matches Longer Subject Heading or Reference?
No
A
Yes
Alphabetical Approach
One Entry Selected?
No
B
Yes
Exact Approach
A
Title-keyword Search
Query Produces Retrievals?
No
C
Yes
Display Titles
C
Query Misspelled?
No
Yes
Correct Spelling
B
Continue Searching?
No
Figure 4.1
Yes
A

Figure 4.3. Search tree for one-word queries

The search tree for one-word queries is given in figure 4.3. One-word queries that matched the initial characters in longer controlled subject headings or *see* references were submitted to the alphabetical approach. The user's selection of a listed controlled vocabulary term invoked the exact approach. Remaining one-word queries were submitted to title-keyword searches. When title-keyword searches failed to produce retrievals, the Blue System presented them to users and asked them to check and correct spelling. If users corrected spelling, the Blue System submitted revised queries to the initial search tree (figure 4.1) because revised queries could have matched exact or normalized forms of controlled vocabulary terms.

Remaining queries for subjects generally were composed of two or more words. Some queries matched the initial words of longer controlled vocabulary terms. These queries were submitted to the alphabetical approach (figure 4.4). The rest were submitted to a series of keyword searches that began with the keyword-in-main-heading search (figures 4.5A–4.5B). When the particular keyword search produced zero or too few retrievals, the Blue System continued searching using the next keyword approach in the series.

Figure 4.4 is a search tree for queries composed of more than one word. The Blue System responded with the alphabetical approach to queries that matched longer controlled vocabulary terms. To find additional material, the system continued searching using the keyword-in-main-heading search. This search tree took into account that some queries were partial matches of controlled vocabulary terms. For example, the query "civil rights movement" matched the first two words and part of the third word in the subject heading "Civil rights movements." The initial system response to such queries was the alphabetical approach. When users responded to system prompts to continue searching using the original query, the Blue System continued with the results of title-keyword searches.

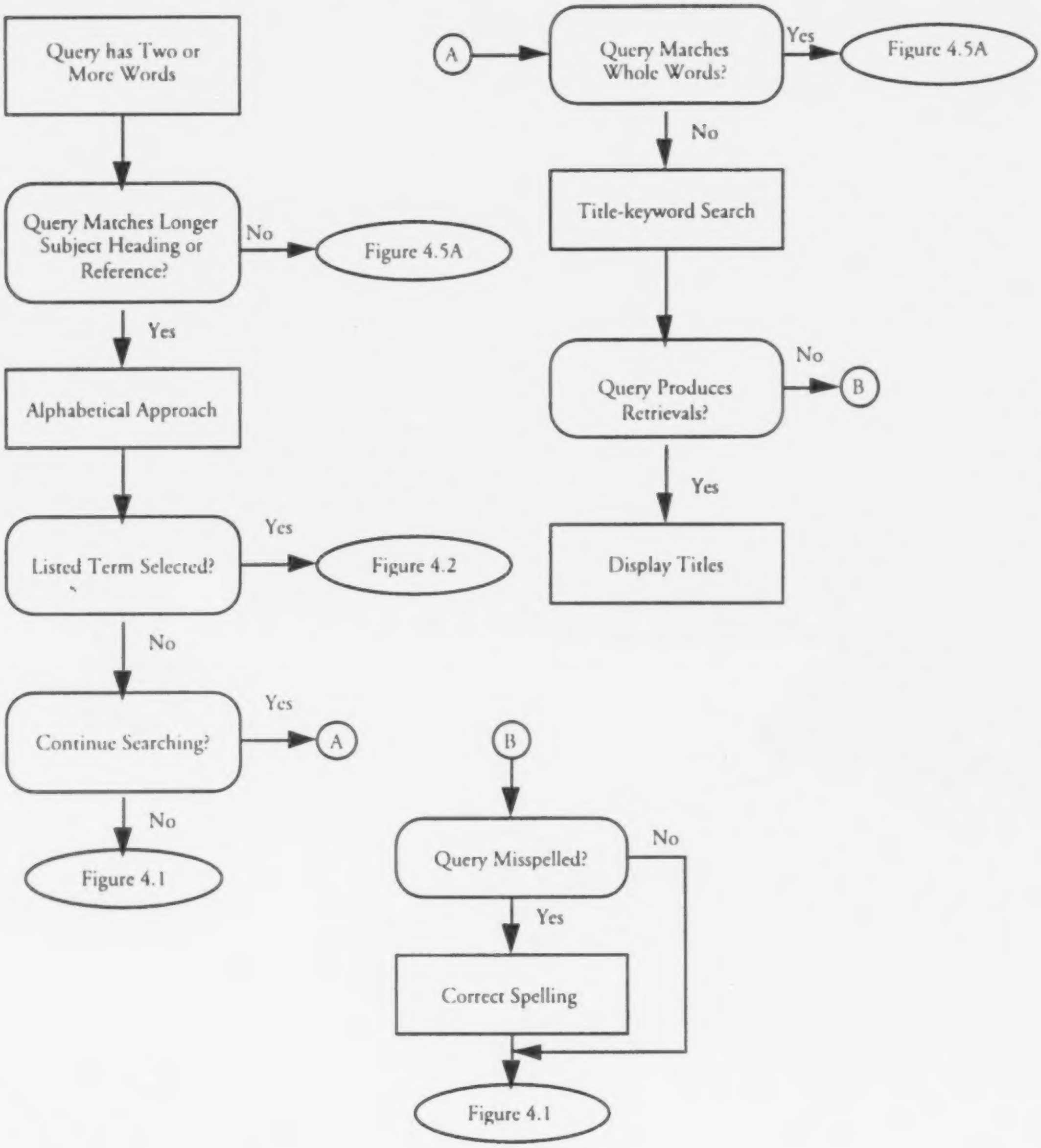

Query has Two or More Words
Query Matches Longer Subject Heading or Reference?
No
Figure 4.5A
Yes
Alphabetical Approach
Listed Term Selected?
Yes
Figure 4.2
No
Continue Searching?
Yes
A
No
Figure 4.1
A
Query Matches Whole Words?
Yes
Figure 4.5A
No
Title-keyword Search
Query Produces Retrievals?
No
B
Yes
Display Titles
B
Query Misspelled?
No
Yes
Correct Spelling
Figure 4.1

Figure 4.4. Search tree for multi-word queries featuring the alphabetical approach

Remaining queries composed of two or more words were submitted to a search tree for keyword approaches (figures 4.5A–4.5B). The search tree shown in figure 4.5A featured the submission of queries to controlled vocabulary searches.

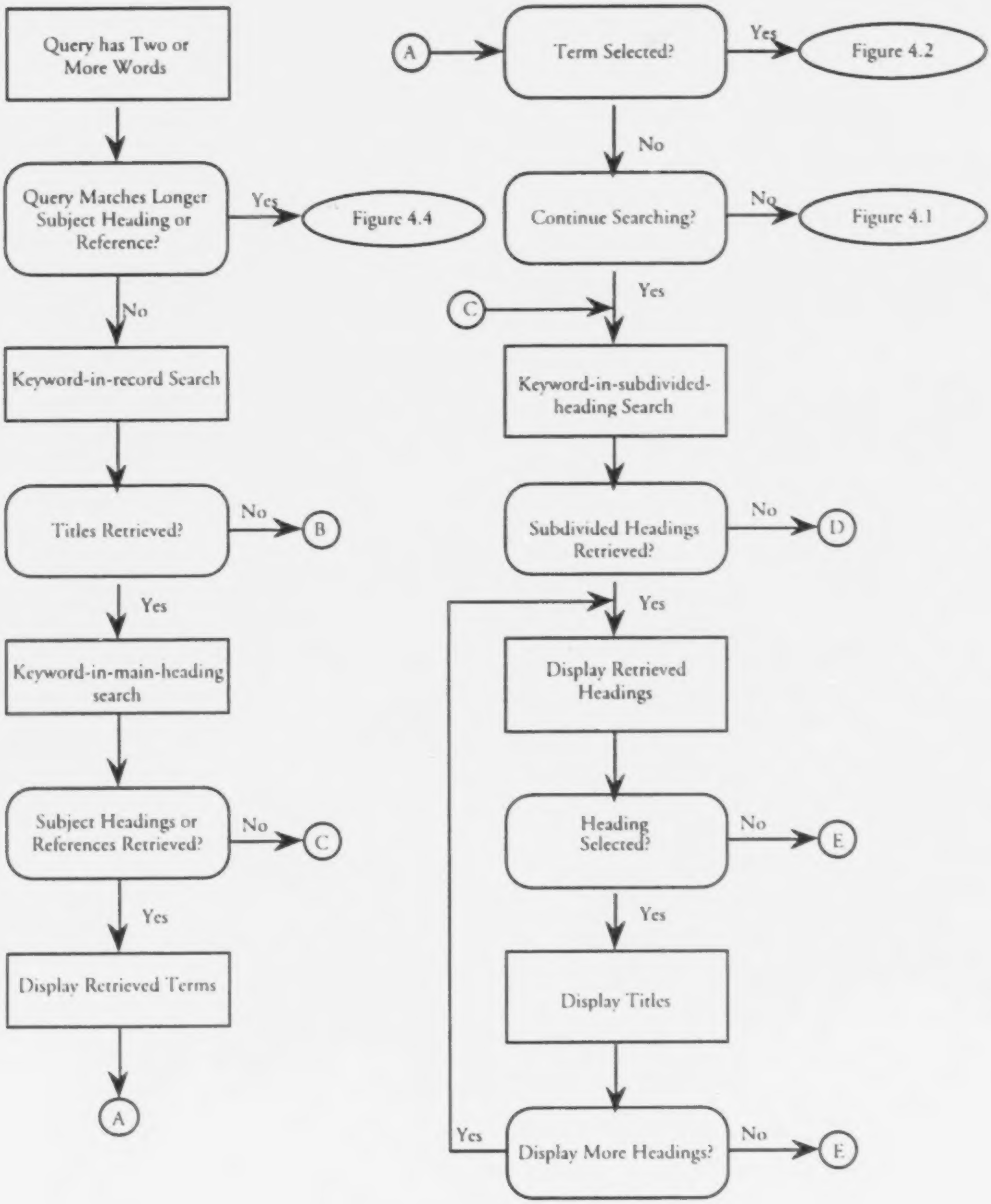

Query has Two or More Words
Query Matches Longer Subject Heading or Reference?
Yes
Figure 4.4
No
Keyword-in-record Search
Titles Retrieved?
No
B
Yes
Keyword-in-main-heading search
Subject Headings or References Retrieved?
No
C
Yes
Display Retrieved Terms
A
A
Term Selected?
Yes
Figure 4.2
No
Continue Searching?
No
Figure 4.1
Yes
C
Keyword-in-subdivided-heading Search
Subdivided Headings Retrieved?
No
D
Yes
Display Retrieved Headings
Heading Selected?
No
E
Yes
Display Titles
Display More Headings?
Yes
No
E

Figure 4.5A. Search tree for multi-word queries featuring keyword approaches

Prior to controlled vocabulary searches, the Blue System first used the keyword-in-record search to ensure that individual query words were posted in the catalog. If one or more query words failed to produce retrievals, the entire series of keyword searches would fail. The Blue System asked users to check the spelling of their queries and correct them if necessary. If users made changes, the Blue System started at the beginning of the subject searching process, i.e., looking for matches of exact and normalized forms of controlled vocabulary terms. If changes were not made, this system asked users for a different query because their original one failed to produce retrievals using any keyword approach. Retrievals produced through the initial keyword-in-record search were not shown to users. Instead, the Blue System continued searching beginning with the keyword-in-main-heading search. The initial keyword-in-record search was meant to save the system additional steps searching for a query that did not retrieve any bibliographic records through the entire series of keyword searches.

Figure 4.5B features free-text searches for multi-word queries. The first free-text search was the title-keyword search. This search was followed by keyword in subject heading fields and keyword-in-record searches.

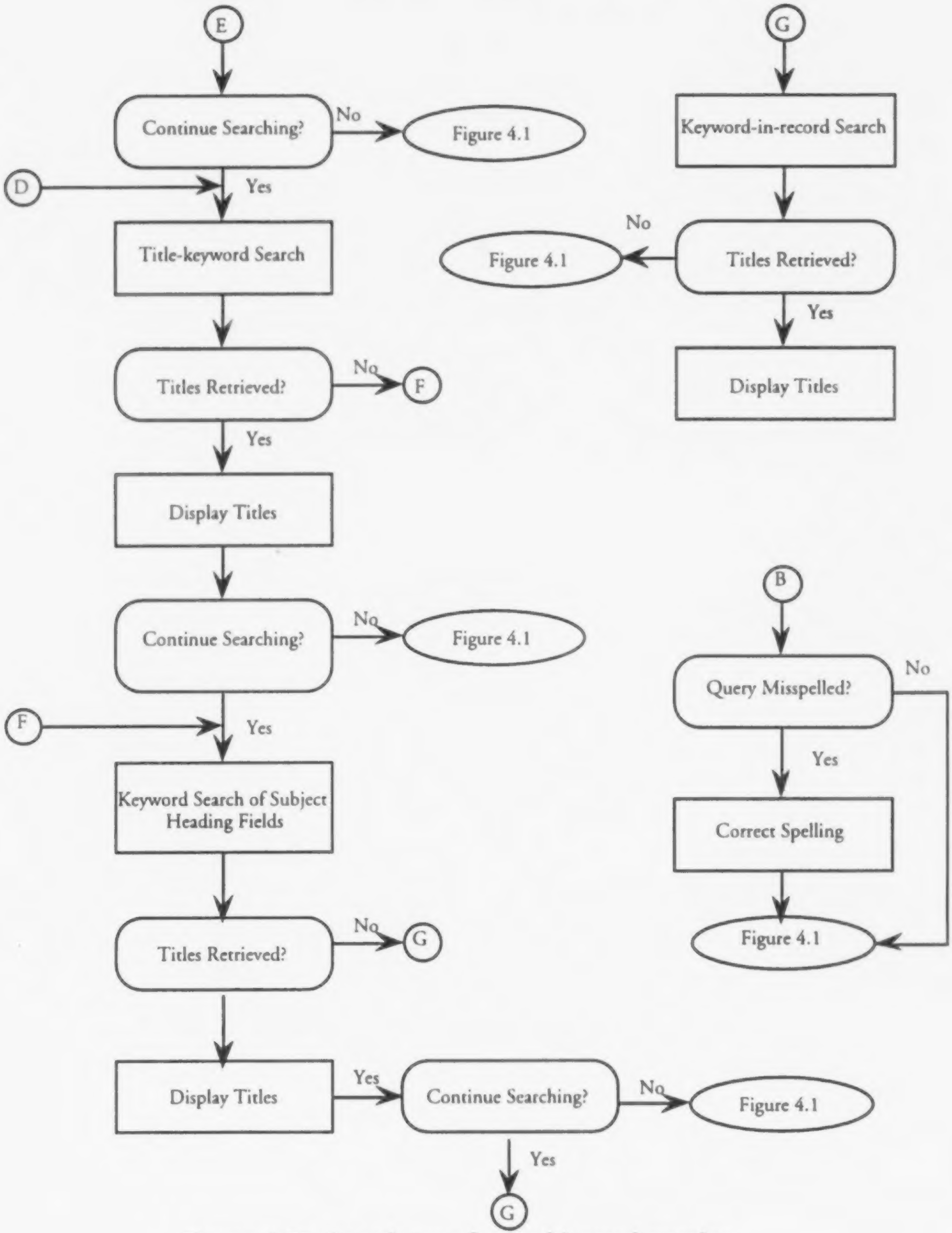

E
Continue Searching?
No
Figure 4.1
D
Yes
Title-keyword Search
Titles Retrieved?
No
F
Yes
Display Titles
Continue Searching?
No
Figure 4.1
F
Yes
Keyword Search of Subject Heading Fields
Titles Retrieved?
No
G
Display Titles
Yes
Continue Searching?
No
Figure 4.1
Yes
G
G
Keyword-in-record Search
No
Figure 4.1
Titles Retrieved?
Yes
Display Titles
B
Query Misspelled?
No
Yes
Correct Spelling
Figure 4.1

Figure 4.5B. Search tree for multi-word queries featuring keyword approaches (contd.)

4.5 Search Tree for Personal-name Queries

Both the Pinstripe and Blue Systems prompted users for surname, given name, and topical elements of their personal-name queries. When searching, the Blue System was selective about the personal-name elements it submitted to various subject searching approaches for personal names. In contrast, the Pinstripe System disregarded its knowledge of the individual elements of personal-name queries and performed keyword-in-record searches for all elements of user queries.

Table 4.1 lists the sequence of Blue-system subject searching approaches to which elements of personal-name queries were submitted. If an approach failed to produce retrievals, the Blue System chose the next approach and element(s) on the list.

Table 4.1. Sequence of personal-name query elements

	Available elements			
Approach	**Last name**	**First name**	**Middle name**	**Topic**
Queries with topics				
Keyword-in-subdivided-heading	X	X		X
Keyword-in-subdivided-heading	X			X
Keyword-in-record	X	X		X
Keyword-in-record	X			X
Queries without or omitting topic elements				
Alphabetical	X	X	X	
Alphabetical	X	X		
Alphabetical	X			

The first step was for the Blue System to single out queries containing topics. These queries were submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If the Blue System failed to find both elements, it omitted the topic element from the query and continued searching through the alphabetical approach. Personal-name queries consisting exclusively of name elements were submitted to the alphabetical approach only.

4.6 Chapter Summary

The search trees that governed the Blue System's selection of subject searching approaches emanated from the findings of an empirical study of user queries.

User queries for subjects generally were controlled by five search trees. The initial search tree dispatched user queries to other search trees that favored certain subject searching approaches based on summary characteristics of user queries (figure 4.1). The search tree for matches of exact and normalized forms of controlled vocabulary terms favored the exact and alphabetical approaches (figure 4.2). One-word queries were given to a search tree that favored alphabetical or title-keyword searches (figure 4.3). Queries composed of two or more words matching longer controlled vocabulary terms were submitted to alphabetical and keyword-in-heading searches (figure 4.4). Remaining queries were controlled by a search tree that submitted them to a series of keyword searches beginning with the keyword-in-record search to detect spelling errors in query words (figures 4.5A and 4.5B).

User queries for personal names were controlled by a single search tree (Table 4.1). Personal-name queries consisting exclusively of name elements were submitted to the alphabetical approach. Queries bearing topic elements were submitted to various keyword searches in the hopes of satisfying both topic and name elements. Failure to produce retrievals resulted in the alphabetical approach.

Search trees preferred two-step subject searching approaches that enlisted the catalog's controlled vocabulary, i.e., exact, alphabetical, and keyword-in-heading searches. If these searches supplied users with appropriate controlled vocabulary terms for expressing their topics of interest, users could refine their searches using related terms, subdivided forms of the matched or selected heading, or controlled vocabulary terms in close alphabetical proximity to the subject queries users enter into online catalogs.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Drabenstott, Karen Markey, and Diane Vizine-Goetz. 1990. "Search trees for subject searching in online catalogs." *Library Hi Tech* 8, 3: 7–20.

Mitev, Nathalie, Gillian Venner, and Stephen Walker. 1985. *Designing an online public access catalog.* London: British Lib. Library & Information Research Report 39.

O'Neill, Edward T., and Rao Aluri. 1979. *Subject heading patterns in OCLC monographic records.* Columbus, OH: OCLC. ERIC ED 183167.

Walker, Stephen, and Richard M. Jones. 1987. *Improving subject retrieval in online catalogues; 1. Stemming, automatic spelling correction and cross-reference tables.* London: British Library. British Library Research Paper no. 24.

5 ASTUTE Subject Searching Functionality

5.1 Introduction

Chapter 5 describes the subject searching functionality of the ASTUTE experimental online catalog. ASTUTE was based on the FoxPro database management system, thus, this chapter begins with a brief description of FoxPro.

ASTUTE was composed of two experimental online catalogs: (1) the Blue System in which search trees controlled the system's selection of a subject searching approach in response to user queries, and (2) the Pinstripe System which chose subject searching approaches randomly in response to user queries. Thus, this chapter's discussion of ASTUTE functionality covers the functionality of both Blue and Pinstripe Systems. It also gives examples of searches in the Blue and Pinstripe Systems to illustrate functionality and system responses to user queries. The discussion occasionally refers to the search trees that control the Blue System's responses that are covered in chapter 4.

5.2 FoxPro Database Management System

The ASTUTE project team sought a relational database management system (RDBMS) to store bibliographic and authority data efficiently on the Gateway microcomputer's hard disk and feature fast retrievals and character manipulation. In relational database management systems (RDBMSs), a database is sometimes called a table. Each separate unit of database information is called a record and each entity within the record is called a field. To feature fast lookups that join two tables or databases, databases can be related to one another by a field that is common to both databases. Databases are usually indexed on this common field. The advantage of joining databases by common fields is to minimize the duplication of information in many records and databases. An example is the bibliographic main record (bib_main) database of the ASTUTE experimental online catalog. Each record in this database had a unique series number and contained information unique to individual items such as the author name(s), Library of Congress classification number, title, imprint, etc. The bibliographic main database was the "parent" database. ASTUTE also had a subject heading database that contained records for subject headings and a series number that joined subject headings to the main database. Since there might be more than one subject heading or subject record for each item, the relationship of the main database to the subject heading database was a one to many relationship and the subject heading database was called the "child."

The project team sought an RDBMS that would retrieve records quickly, have window manipulation, low-level file access functions, ease of screen manipulation, and be a full-featured programming language. The primary consideration was the speed of retrieval. The many character string searches utilized by the program would make the speed of record retrieval the most important issue in the selection of the RDBMS. After considering several options, the ASTUTE project team chose the FoxPro 2.0 relational database management system. FoxPro 2.0 included many built-in functions that manipulated character strings needed to build databases and access them quickly. FoxPro could also store variable-length records of any size in a field type called a "memo" field. Built-in tools in FoxPro included access to a browse capability that was very useful in allowing users to scroll through lists and choose a topic from them. Another useful function was the ease of building screens that contained objects or buttons for the user to select. The built-in screen functions of FoxPro 2.0 also allowed users to use the keyboard or the mouse to select an objects on the screen.

5.3 Exact approach to Subject Searching

5.3.1 Background

The exact approach to subject searching is a *new* subject searching approach. This approach *facilitates browsing* for users who have entered or selected a posted subject heading in an online catalog.

The exact approach was designed in response to heavy criticism of the alphabetical approach to subject searching (Drabenstott and Vizine-Goetz, 1994, 241). Recent studies of online catalog searching have demonstrated that the alphabetical approach is no longer capable of managing large numbers of subdivided forms of subject headings because users demonstrate little perseverance when faced with large numbers of retrievals (Van Pulis and Ludy 1988; Wiberley and Daugherty 1988). Several researchers have recognized the difficulty of displaying long lists of subdivided forms of headings to users and suggested methods for reducing such lists (Holley and

Killheffer 1982, 130; Cochrane 1986, 62; Markey 1984, 112–3; and Chan 1986, 355–8; Massicotte 1988, 374).

The exact approach gives users new opportunities for browsing. It anticipates the user's selection of the exact match from an alphabetical list, and thus begins with a report of the results of such matches. The report includes a summary of subdivided forms of the matched subject heading in the form of broad categories, and, if available, options for browsing related terms and other information about the matched heading. The challenge here is to convey to users the range of available opportunities in a single screen.

Implementations of the exact approach in existing online catalogs are limited in usefulness because only three broad categories can be constructed from the three different types of subject subdivisions used with subject headings, viz. topical subdivisions ($x), period subdivisions ($y), and geographic subdivisions ($z). Defining new broad categories would be a huge undertaking because it could require editorial review of the thousands of subdivisions in SCM:SH. Although such a review is not within the scope of this project, the ASTUTE project team was able to expand the number of broad categories from three to four by reviewing topical subdivisions, singling out those that describe form, and assigning them a unique subfield code ($1) in bibliographic records. (See section 3.3.4.1 for details on this effort.) Of the many broad categories that they could have created, the ASTUTE project team created the form category because of recommendations from the LC Subject Subdivisions Conference that called for the creation of form subdivisions (Conway 1992, 9).

5.3.2 Generating Entries for the Exact Approach Database

In the ASTUTE experimental online catalog, computer programs generated normalized forms of subject headings in bibliographic records. They also generated normalized forms of *see* references for the subject headings used in bibliographic records. The programs added normalized forms of subject headings and *see* references into the exact_approach database. When users entered their query, ASTUTE searched the exact_approach database to find a match. Normalized forms of subject headings and *see* references were generated by manipulating headings and references in the following ways:

- Converting upper to lower case.
- Eliminating punctuation (i.e., commas, parentheses, apostrophes, hyphens, semicolons, periods, slashes, subfield codes).
- Eliminating stopwords (i.e., the, of, and, or, in, etc., for, to).

- Eliminating qualifiers.
- Normalizing word order by alphabetizing words in subject headings.
- Double posting subject headings bearing hyphenated words (without hyphens and as single run-on words) and bearing apostrophes (with and without trailing "apostrophe-s").

Examples of normalized forms of subject headings are listed below adjacent the authorized form from which they were generated:

Subject heading	**Normalized form**
Afro-American children	afro american children, afroamerican children
Astronomy, Prehistoric	astronomy prehistoric
Ball's Bluff, Battle of, 1861	1861 balls battle bluff, 1861 ball battle bluff
Computers $x Acronyms	acronyms computers
Expert systems (Computer Science)	computer expert science systems, expert systems

Table 5.1 lists fields and subfields of USMARC bibliographic and subject authority records that contributed subject headings to ASTUTE's exact approach database. Since the exact approach was limited to user queries for subjects generally, fields and subfields for personal names as subjects did not contribute to the exact approach database.

Table 5.1. Fields/Subfields in Exact Search Approach Database

USMARC Field Name	Format Type	Tag	Subfields
Subject (corporate body)	bibliographic	610	abndcxyztps
Subject (meeting name)	bibliographic	611	andcktxyz
Subject (uniform title)	bibliographic	630	apsdxyztnk
Subject (topical)	bibliographic	650	axyz
Subject (geographic)	bibliographic	651	axyz
See reference (topical)	authority	450	axyz
See reference (geographic)	authority	451	axyz

5.3.3 Responding to User Queries with the Exact Approach

The exact approach was limited to the Blue System. This system responded with the exact approach to user queries for subjects generally that matched normalized forms of subject headings from bibliographic records and/or *see* references from subject authority records. The exact approach offered users many options to further the subject search. Table 5.2 lists these options in six option boxes numbered 5.2.1 to 5.2.6. The wording of options in Table 5.2 generally corresponds to the wording on ASTUTE screens. Searchers pursued options by positioning the cursor over highlighted text using key combinations, e.g., directional and "Enter" keys, <Tab> and <Enter> keys, or moving and clicking the mouse. ASTUTE did not require users to explicitly type in option names. For example, users pursued backward and forward browsing options using a mouse to manipulate a scroll bar.

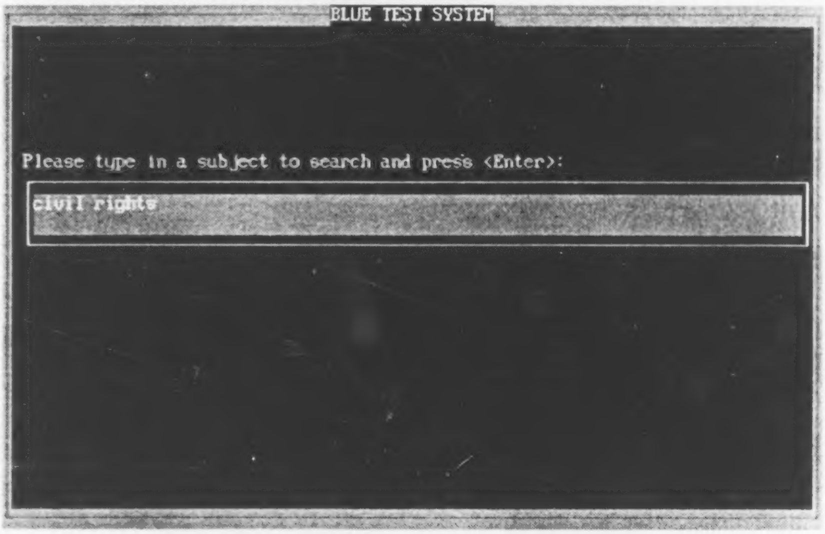
The Blue System responded to user queries that matched a normalized form(s) of subject heading or *see* reference with the "Exact approach main menu" (options 5.2.2, Table 5.2). When the Blue System found more than one matching subject heading and/or *see* reference, it presented users with an intermediary screen bearing matched heading(s) and/or reference(s) and prompted them to choose a listed term (options 5.2.1, Table 5.2). User selection of a listed term resulted in the exact approach main menu (options 5.2.2, Table 5.2).

Table 5.2. Exact Search Approach Options

Intermediary term selection options 5.2.1:	
[Intermediary matched terms list]	
Choose listed heading	==> Options 5.2.2
Main menu display options 5.2.2:	
[Exact approach main menu]	
Display general works	==> Options 5.2.3
Display narrower terms	==> Options 5.2.4
Display broader terms	==> Options 5.2.4
Display notes on this subject	==> Options 5.2.5
Display specialized subtopics	==> Options 5.2.6
Display specific places	==> Options 5.2.6
Display various time periods	==> Options 5.2.6
Display forms of publication	==> Options 5.2.6
Go to previous screen	==> Options 5.2.2*
Expand search	==> Search trees
Start over	==> Switch systems
Bibliographic Record Display Options 5.2.3:	
[Bibliographic record display]	
Display next record	==> Options 5.2.3
Display previous record	==> Options 5.2.3

Display second page of record	==> Options 5.2.3
Exit record display	==> Options 5.2.2
Related Terms Display 5.2.4:	
[Broader or narrower terms display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected term	==> Options 5.2.2
Go back	==> Options 5.2.2
Scope Note Display Options 5.2.5:	
[Scope note display]	
Browse backward or forward	==> Manipulate scroll bar
Go back	==> Options 5.2.2
Subtopics Display Options 5.2.6:	
[Subtopics display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected subtopic	==> Options 5.2.2 or 5.2.6
Go back	==> Options 5.2.2

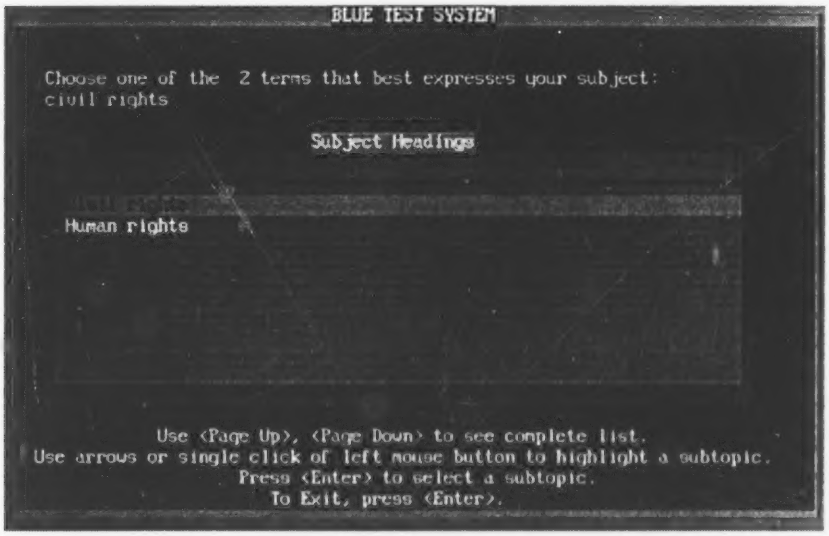
*Option appears on subsequent exact approach main menus.



```
Please type in a subject to search and press <Enter>:
civil rights
```

Figure 5.1. User entry of "civil rights" query

BEST COPY AVAILABLE



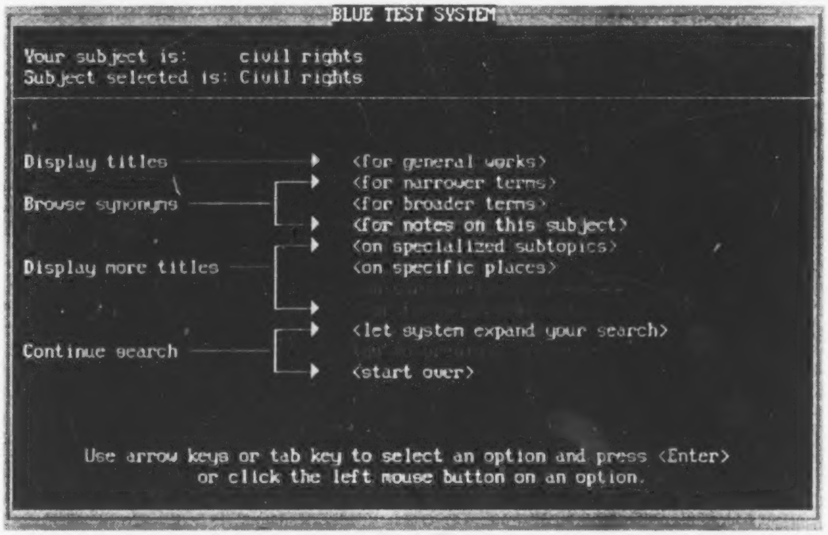
```
                        BLUE TEST SYSTEM

Choose one of the  2 terms that best expresses your subject:
civil rights

                        Subject Headings

 Civil rights
 Human rights



          Use <Page Up>, <Page Down> to see complete list.
Use arrows or single click of left mouse button to highlight a subtopic.
                 Press <Enter> to select a subtopic.
                      To Exit, press <Enter>.
```

Figure 5.2. Intermediary matched terms list for "civil rights"

Figures 5.1–5.8 show the Blue System's response to the user query "civil rights." In figure 5.1, the user enters the query "civil rights" into the Blue System. The system's response to this query is an intermediary screen prompting users to choose between two terms, "Civil rights" and "Human rights." The Blue System gives users a choice of terms because the query matches the exact form of the subject heading "Civil rights" and the normalized form of the *see* reference "Civil rights (International law)" under which is the authorized subject heading "Human rights" (options 5.2.1, Table 5.2)

The user's response to the intermediary screen of matched terms is to select the listed term "Civil rights" (figure 5.2). The Blue System responds with the exact approach main menu (options 5.2.2, Table 5.2, and figure 5.3). Although the exact approach main menu gives users ten options, there are many searches in which one or more options are disabled. In figure 5.3, the Blue System constructs a main menu in response to the user query "civil rights." Options for displaying titles on "various time periods" and "forms of publication" and for continuing the search by going to the "previous screen" are disabled.

BEST COPY AVAILABLE

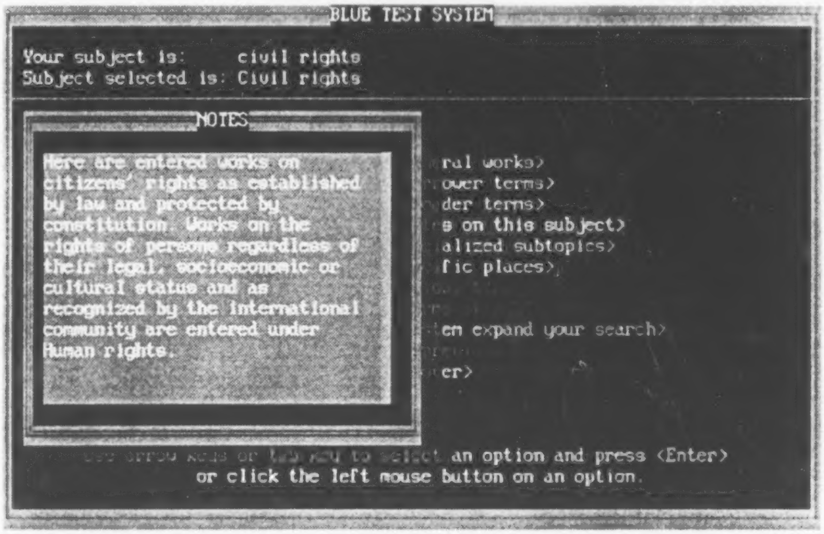

BLUE TEST SYSTEM
Display titles
Browse synonyms
Display more titles
Continue search
<for general works>
<for narrower terms>
<for broader terms>
<for notes on this subject>
<on specialized subtopics>
<on specific places>
<let system expand your search>
<start over>

Your subject is: civil rights
Subject selected is: Civil rights

Use arrow keys or tab key to select an option and press <Enter>
or click the left mouse button on an option.

Figure 5.3. Exact approach main menu for "Civil rights"

User selection of the "notes on this subject" option results in a display of the LCSH scope note for the subject heading "Civil rights" (figure 5.4). Scroll bars are enabled for long scope notes to allow users to scan forward and backward. When users finish displaying scope notes, they return to the main menu (options 5.2.2, Table 5.2).


BLUE TEST SYSTEM
NOTES
ral works>
ower terms>
der terms>
s on this subject>
alized subtopics>
fic places>
em expand your search>
er>

Your subject is: civil rights
Subject selected is: Civil rights

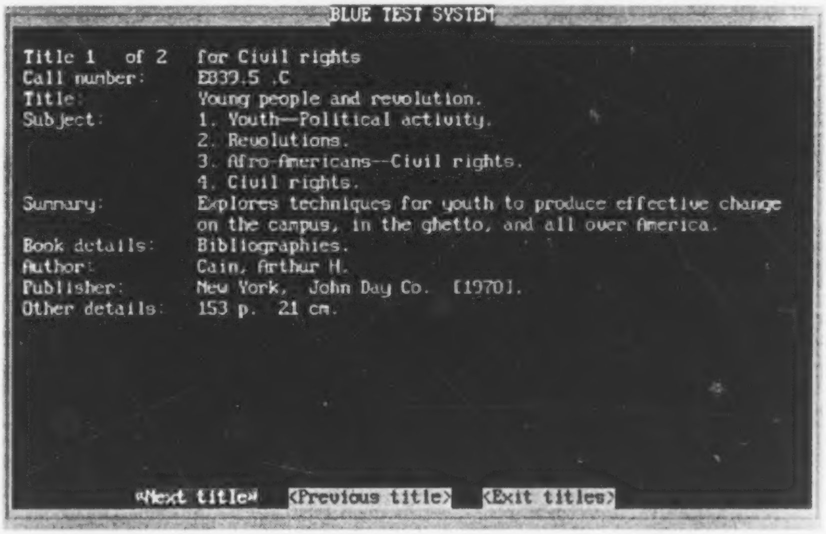
Here are entered works on citizens' rights as established by law and protected by constitution. Works on the rights of persons regardless of their legal, socioeconomic or cultural status and as recognized by the international community are entered under Human rights.

Use arrow keys or tab key to select an option and press <Enter>
or click the left mouse button on an option.

Figure 5.4. Scope note for "Civil rights"

BEST COPY AVAILABLE

At the main menu, user selection of the "general works" option results in a display of bibliographic records bearing the subject heading "Civil rights" (figure 5.5).

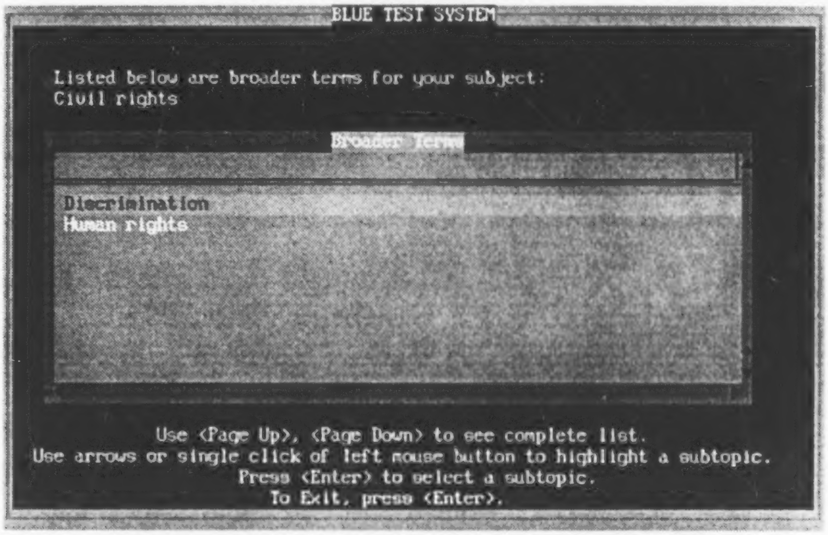


```
Title 1   of 2    for Civil rights
Call number:      E839.5 .C
Title:            Young people and revolution.
Subject:          1. Youth--Political activity.
                  2. Revolutions.
                  3. Afro-Americans--Civil rights.
                  4. Civil rights.
Summary:          Explores techniques for youth to produce effective change
                  on the campus, in the ghetto, and all over America.
Book details:     Bibliographies.
Author:           Cain, Arthur H.
Publisher:        New York,  John Day Co.  [1970].
Other details:    153 p.  21 cm.
```

Figure 5.5. Bibliographic record bearing subject heading "Civil rights"

When displaying bibliographic records, users have options to display the next record, previous record, exit records, or, display the second page of a lengthy record (options 5.2.3, Table 5.2). In response to a user selection of the exit records option, the Blue System returns users to the exact approach main menu (options 5.2.2, Table 5.2, and figure 5.3). Users could continue to explore by choosing options to display narrower or broader terms, display titles on specialized subtopics or specific places, or let the system expand their search. For example, user selection of the broader term option produces the display of two subject headings shown in figure 5.6. If the user selects a listed broader term, the Blue System responds with an entirely new exact approach main menu bearing a different set of enabled and disabled options from the exact search main menu (options 5.2.2, Table 5.2). For example, the user's selection of the broader term "Discrimination" in figure 5.6 produces the new exact search main menu in figure 5.7 that includes enabled options for general works, broader and narrower terms, notes, specific places, expanding the search, starting over, and going to the previous screen.

BEST COPY AVAILABLE

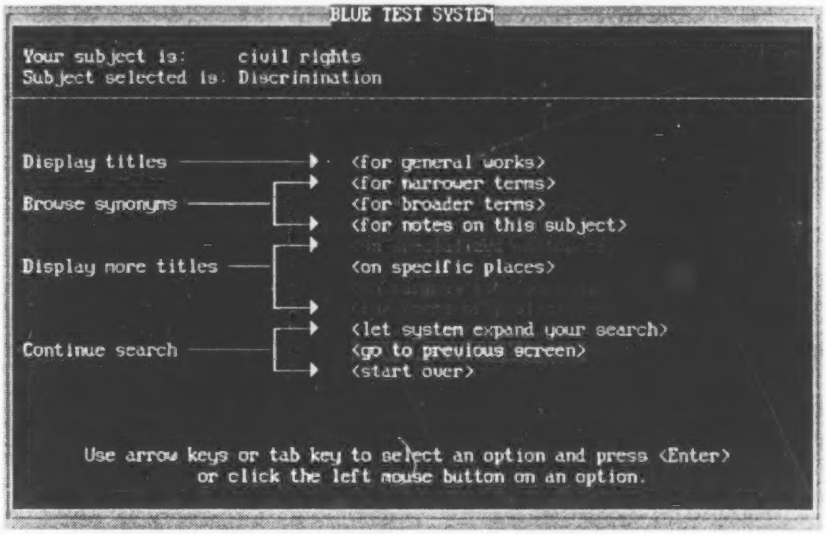

BLUE TEST SYSTEM
Broader Terms

Listed below are broader terms for your subject:
Civil rights

Discrimination
Human rights

Use <Page Up>, <Page Down> to see complete list.
Use arrows or single click of left mouse button to highlight a subtopic.
Press <Enter> to select a subtopic.
To Exit, press <Enter>.

Figure 5.6. Broader terms for "Civil rights"


BLUE TEST SYSTEM
Display titles
<for general works>
<for narrower terms>
Browse synonyms
<for broader terms>
<for notes on this subject>
Display more titles
<on specific places>
<let system expand your search>
Continue search
<go to previous screen>
<start over>

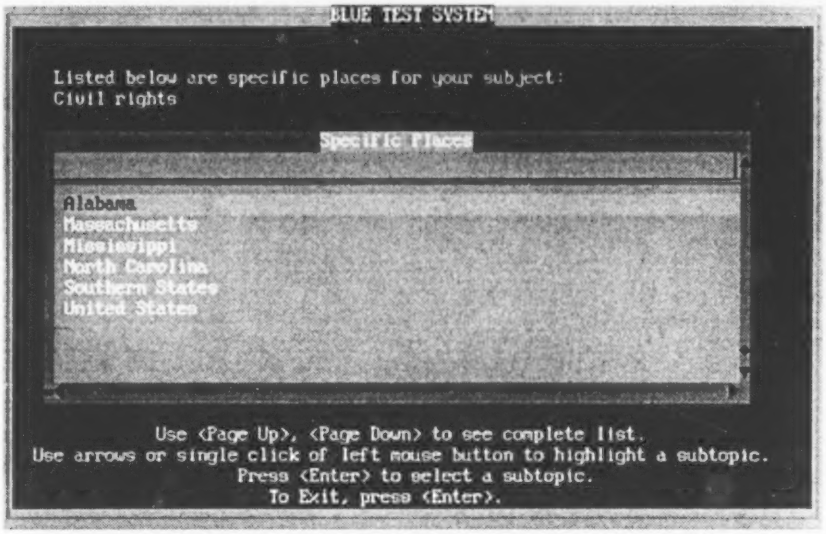
Your subject is: civil rights
Subject selected is: Discrimination

Use arrow keys or tab key to select an option and press <Enter>
or click the left mouse button on an option.

Figure 5.7. New exact search main menu for "Discrimination"

In figure 5.7, the user's selection of the previous screen option produces the initial main menu. That is, the user's selection of the "previous screen" option in the new main menu shown in figure 5.7 produces the original main menu shown in figure 5.3.

BEST COPY AVAILABLE



Listed below are specific places for your subject:
Civil rights

Alabama
Massachusetts
Mississippi
North Carolina
Southern States
United States

Use <Page Up>, <Page Down> to see complete list.
Use arrows or single click of left mouse button to highlight a subtopic.
Press <Enter> to select a subtopic.
To Exit, press <Enter>.

Figure 5.8. Place subtopics under "Civil rights"

Users could continue exploring by choosing other enabled options such as the option to display specific places. Figure 5.8 shows a list of places produced in response to the user selection of this option. Users manipulate the scroll bar to browse backward and forward in the list (options 5.2.6, Table 5.2). User selection of a listed place results in an entirely new exact approach main menu (options 5.2.2, Table 5.2).

When users select the "Expand search" option, the search trees take control of the Blue System and execute a search for the original query using the next search approach on the search trees that produces retrievals (options 5.2.2, Table 5.1). For example, in response to the "Expand search" option selected by the user interested in "civil rights," the Blue System executes an alphabetical search because it is the next subject search approach in the search tree for multi-word queries and the user query "civil rights" matches longer subject headings, e.g., "Civil rights demonstrations," "Civil rights movements."

The Blue System responds to users selecting the "Start over" option in the exact approach main menu by switching to the Pinstripe System or beginning the first of eleven post-search questions (options 5.2.2, Table 5.2).

BEST COPY AVAILABLE

5.4 Alphabetical Approach to Subject Searching

5.4.1 Background

The alphabetical approach to subject searching is recommended for user queries that match longer subject headings or *see* references. Systems respond with an alphabetical list of subject headings and *see* references in the neighborhood of the matched term. The objective of displaying alphabetical lists of controlled vocabulary terms to users is to spark their interest in a term that is in close alphabetical proximity to their query. When users select a term from the alphabetical list, systems respond to their selection with the exact approach. The exact approach gives users access to the unique subdivided forms of matched headings, related terms, and other information about the matched heading from LCSH-mr.

When there are many subdivided forms connected with subject headings, the alphabetical approach is not effective for displaying the unique subdivided forms of assigned subject headings because users demonstrate little perseverance when faced with large numbers of retrievals (Van Pulis and Ludy 1988). The alphabetical approach should be treated as an opportunity to stimulate users with controlled vocabulary when their queries do not quite match specific terms. Thus, subject headings displayed to users in an alphabetical approach should be limited to unsubdivided subject headings.

Including subdivided headings in the alphabetical index may induce users to narrow their topics based on the subdivided forms displayed to them. Access to subdivided forms of matched headings are accessible to users who choose a listed subject heading through the exact approach. In the alphabetical approach, the only subdivided forms that systems should display to users are subdivided *see* references.

5.4.2 Responding to User Queries with the Alphabetical Approach

The alphabetical approach was implemented slightly differently in the Blue and Pinstripe Test Systems. When user queries matched longer subject headings or references, the search trees required the Blue System to respond with the alphabetical approach. When the algorithm that governed the Pinstripe system's selection of a subject searching approach selected the alphabetical approach, the Pinstripe System responded with the alphabetical approach regardless of the extent to which the user query matched subject headings or *see* references.

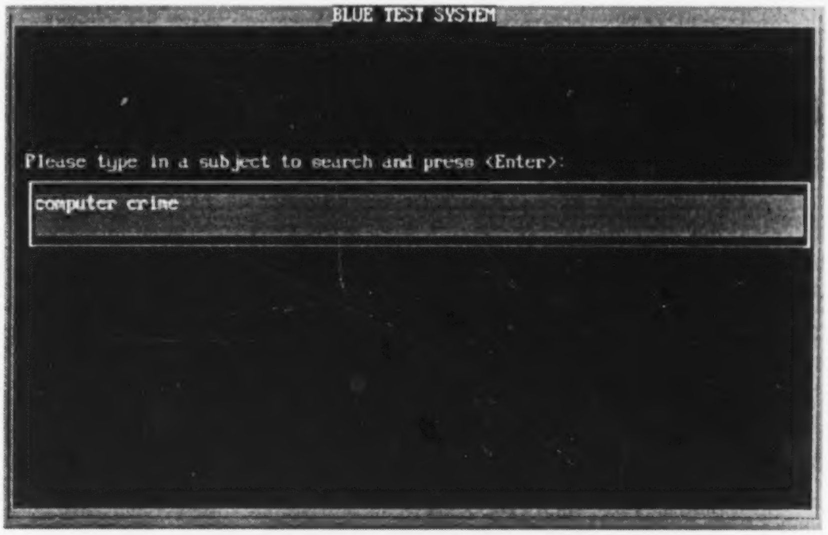
The two test systems responded to user queries matching a longer subject heading or *see* reference with an alphabetical list of subject headings in the alphabetical neighborhood of the partially matched or entered query. Thus, the alphabetical approach was merely a prelude or preface to the exact approach. User selection of a listed term invoked the exact approach. All options available to users in the exact approach (Table 5.2) were also available to them upon the selection of a listed heading or reference in the alphabetical approach. Table 5.3 lists these options in seven option boxes numbered 5.3.0 to 5.3.6. Options for alphabetical term selection (options 5.3.0) were not included among the exact search approach options. One option works totally different from its counterpart in the exact search: the "go to previous screen" option returns users to the alphabetical list of subject heading and references.

Table 5.3. Alphabetical Search Approach Options

Alphabetical term selection options 5.3.0:	
[Alphabetical term list]	
Choose listed heading	==> Options 5.3.1 or 5.3.2
Browse forward or backward in list	==> Click on scroll bar
Let system expand search	==> Search trees
Undo	==> Options 5.3.0
Start over	==> Switch systems
Intermediary term selection options 5.3.1:	
[Intermediary matched terms list]	
Choose listed heading	==> Options 5.3.2
Main menu display options 5.3.2:	
[Initial exact approach main menu]	
Display general works	==> Options 5.3.3
Display narrower terms	==> Options 5.3.4
Display broader terms	==> Options 5.3.4
Display notes on this subject	==> Options 5.3.5
Display specialized subtopics	==> Options 5.3.6
Display specific places	==> Options 5.3.6
Display various time periods	==> Options 5.3.6
Display forms of publication	==> Options 5.3.6
Go to previous screen	==> Options 5.3.0
Expand search	==> Search trees
Start over	==> Switch systems

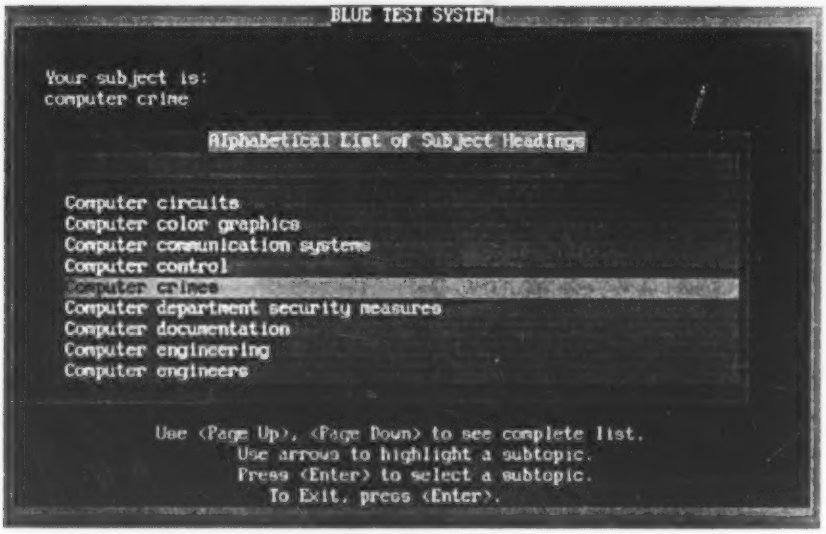
Bibliographic Record Display Options 5.3.3:	
[Bibliographic record display]	
Display next record	==> Options 5.3.3
Display previous record	==> Options 5.3.3
Display second page of record	==> Options 5.3.3
Exit record display	==> Options 5.3.2
Related Terms Display 5.3.4:	
[Broader or narrower terms display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected term	==> Options 5.3.2
Go back	==> Options 5.3.2
Scope Note Display Options 5.3.5:	
[Scope note display]	
Browse backward or forward	==> Manipulate scroll bar
Go back	==> Options 5.3.2
Subtopics Display Options 5.3.6:	
[Subtopics display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected subtopic	==> Options 5.3.2 or 5.3.6
Go back	==> Options 5.3.2

Figures 5.9–5.13 show the Blue System's response to the user query "computer crime." In figure 5.9, the user enters the query "computer crime" into the Blue System. The system's response to this query is an alphabetical list of subject headings bearing the longer term "Computer crimes" and four terms preceding and following this term in the alphabet (figure 5.10).



```
BLUE TEST SYSTEM

Please type in a subject to search and press <Enter>:
computer crime
```

Figure 5.9. User entry of "computer crime" query



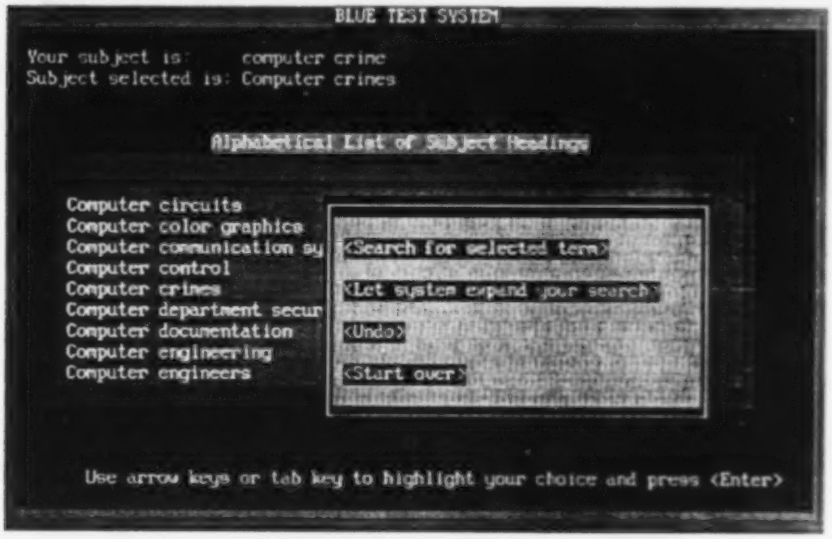
```
BLUE TEST SYSTEM

Your subject is:
computer crime

        Alphabetical List of Subject Headings

  Computer circuits
  Computer color graphics
  Computer communication systems
  Computer control
  Computer crimes
  Computer department security measures
  Computer documentation
  Computer engineering
  Computer engineers

       Use <Page Up>, <Page Down> to see complete list.
             Use arrows to highlight a subtopic.
             Press <Enter> to select a subtopic.
               To Exit, press <Enter>.
```

Figure 5.10. Alphabetical list of subject headings provoked by "computer crime" query

The user selects the listed term "Computer crimes" (figure 5.11). This action results in the exact approach main menu (figure 5.12). This menu gives the user options to display general works, display broader or narrower terms, let the system expand the search, go to previous screen, or start over. If the user selects the "go to previous

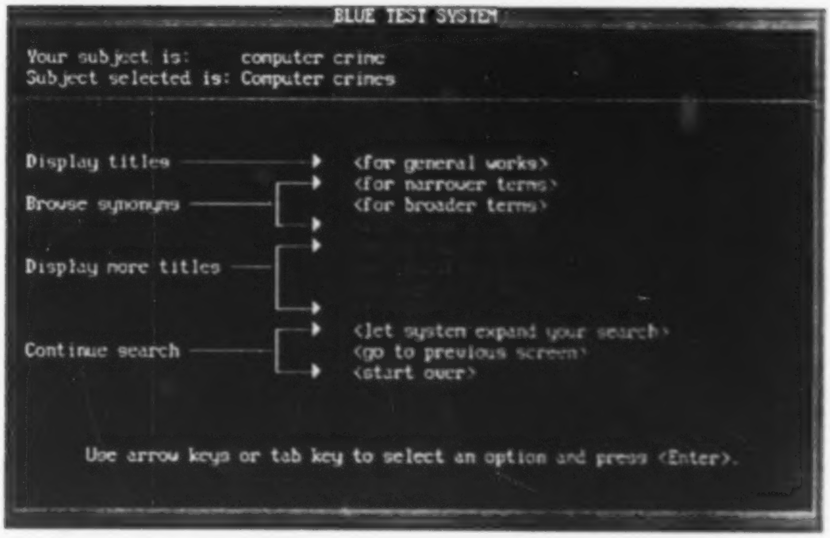
BEST COPY AVAILABLE

screen" option, the Blue System returns to the original alphabetical list of subject headings (figure 5.10).



```
                        BLUE TEST SYSTEM
Your subject is:      computer crime
Subject selected is: Computer crimes

              Alphabetical List of Subject Headings

   Computer circuits
   Computer color graphics
   Computer communication sy  <Search for selected term>
   Computer control
   Computer crimes            <Let system expand your search>
   Computer department secur
   Computer documentation     <Undo>
   Computer engineering
   Computer engineers         <Start over>

     Use arrow keys or tab key to highlight your choice and press <Enter>
```

Figure 5.11. User selection of "Computer crimes"

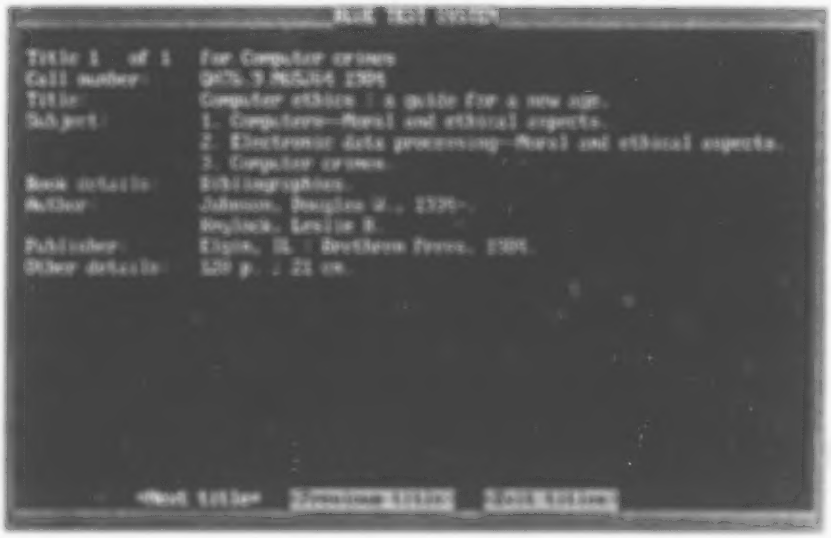


```
                        BLUE TEST SYSTEM
Your subject is:      computer crime
Subject selected is: Computer crimes

Display titles ──────────────►   <for general works>
                        ┌────►   <for narrower terms>
Browse synonyms ────────┤        <for broader terms>
                        └────►
                        ┌────►
Display more titles ────┤
                        └────►
                        ┌────►   <let system expand your search>
Continue search ────────┤        <go to previous screen>
                        └────►   <start over>

     Use arrow keys or tab key to select an option and press <Enter>.
```

Figure 5.12. Exact approach main menu for "Computer crime"

BEST COPY AVAILABLE

The user's selection of the "general works" option retrieves only one title (figure 5.13). To find additional material, the user could explore broader or narrower terms. The user could also choose the "let system expand search" option. This action would induce the search trees to take control of the Blue System and execute a search for the original query (i.e., "computer crime") using the next search approach on the search trees that produces retrievals. The Blue System responds to users selecting the "Start over" option by switching to the Pinstripe System or beginning the first of eleven post-search questions (options 5.3.2, Table 5.3).



```
Title 1   of 1   For Computer crimes
Call number:     QA76.9.M65J64 1984
Title:           Computer ethics : a guide for a new age.
Subject:         1. Computers--Moral and ethical aspects.
                 2. Electronic data processing--Moral and ethical aspects.
                 3. Computer crimes.
Book details:    Bibliographies.
Author:          Johnson, Douglas W., 1934-.
                 [illegible], Leslie R.
Publisher:       Elgin, IL : Brethren Press, 1984.
Other details:   120 p. ; 21 cm.
```

Figure 5.13. Bibliographic record bearing subject heading "Computer crimes"

5.5 Keyword-in-main-heading Search

5.5.1 Background

The keyword-in-main-heading approach was unique to the Blue Test System. There were two situations in which the Blue System invoked keyword-in-main-heading searches: (1) when user queries exceeded one word and failed to produce retrievals through the exact and alphabetical approaches, or (2) when user queries exceeded one word, produced retrievals through the exact and/or alphabetical approaches, but users wished to continue searching. Neither situation guaranteed that keyword-in-main-heading searches would produce retrievals.

When the keyword-in-main-heading approach yielded retrievals, the Blue System responded with an alphabetical list of subject headings and *see* references bearing user queries words in listed main subject headings. The objective of displaying the alphabetical keyword list of controlled vocabulary terms was the same as the objective of displaying an alphabetical list in the alphabetical approach, viz. to spark users' interest in a term that was partially but not entirely composed of words in the user query. When users selected a term from the alphabetical keyword list, the system responded with the exact approach. The exact approach gave users access to the unique subdivided forms of matched headings, related terms, and other information about the matched heading from LCSH-mr.

When user queries matched words embedded in subject headings, the results of keyword-in-main-heading searches were particularly useful in providing relevant searching terminology. For example, the user entering "jazz music" would retrieve an alphabetical list of subject headings bearing words in the user query:

- Double-bass and piano **music** (**Jazz**)
- Guitar **music** (**Jazz**)
- Monologues with **music** (**Jazz** ensemble with chamber orchestra)
- Piano **music** (**Jazz**)
- Saxophone **music** (**Jazz**)

In an alphabetical search, none of these subject headings would be in the alphabetical neighborhood of the subject heading "Jazz (Music)." They would also not be accessible through the exact approach because they are not broader or narrower terms under "Jazz (Music)." Thus, the keyword-in-main-heading approach had the potential to retrieve useful searching terminology that other controlled vocabulary approaches would miss.

5.5.2 Generating Entries for the Keyword-in-main-heading Search

The ASTUTE project team wrote a program to identify words in selected fields of bibliographic records and save them in a database supporting the various keyword approaches of the Blue and Pinstripe Systems. The algorithm for each keyword approach was the same. Thus, procedures used to identify and save words to the database supporting the keyword-in-main-heading database were the same for all other keyword approaches of the Blue and Pinstripe Systems. The program first identified words, converted them to lower case, and eliminated punctuation and characters from the field to be added to the database supporting the keyword-in-main-heading approach. The program eliminated the following punctuation and characters and replaced them with a space and it replaced 's and 's with an s.

etc.
,
.
:
;
(
)
's (replaced with s)
s' (replaced with s)
[
]
*
~
=
$
`
#
ASCII character 27
&

The program then divided phrases into words. Any words separated by a hyphen (-) or slash (/) were also added as separate words. Any slash that was not part of a word was eliminated. The program then eliminated the following words from the word list:

of
and
or
in
the
for
to
by
al
with
at
on
a

The keyword database supporting the keyword-in-main-heading search contained words from only the $a subfield of 650 fields of bibliographic records.

5.5.3 Responding to User Queries with the Keyword-in-main-heading Search

The keyword-in-main-heading approach was only implemented in the Blue Test System. This approach was the third approach (after exact and alphabetical approaches) that the Blue System enlisted to produce retrievals for user queries exceeding one word. When user queries matched words in main subject headings, the Blue System responded with an alphabetical keyword list of matched headings. User selection of a listed term invoked the exact approach. Thus, the keyword-in-main-heading approach (like the alphabetical approach) was merely a prelude or preface to the exact approach.

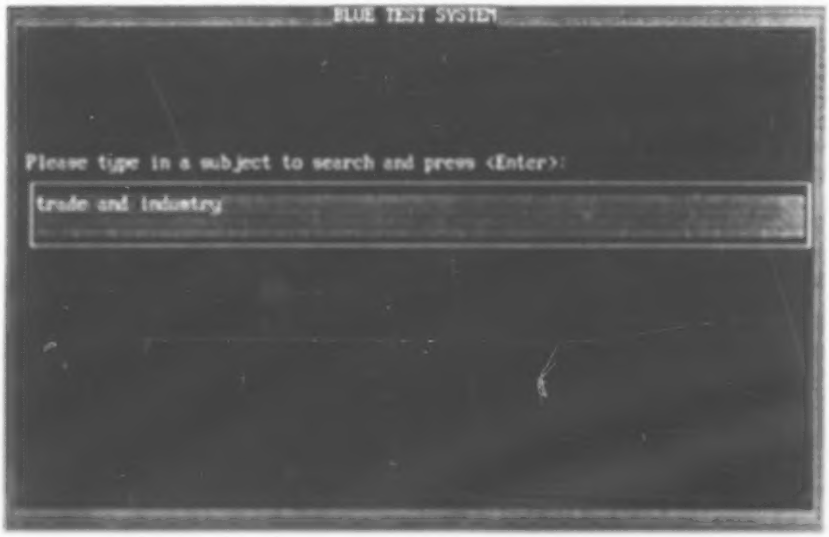
All options available to users in the exact approach (Table 5.2) were also available to them upon the selection of a listed heading in the keyword-in-main-heading approach. Table 5.4 lists these options in six option boxes numbered 5.4.0 to 5.4.5. Options for the selection of a term from the alphabetical keyword list (options 5.4.0) were not included in exact search approach options (Table 5.2). One option worked totally different from its counterpart in the exact search: the "go to previous screen" option returned users to the alphabetical keyword list of subject headings. The wording of options in Table 5.4 generally corresponds to the wording on system screens.

Table 5.4. Keyword-in-main-heading Search Approach Options

Alphabetical keyword list options 5.4.0:	
[Alphabetical term list]	
Choose listed heading	==> Options 5.4.1
Let system expand search	==> Search trees
Undo	==> Options 5.4.0
Start over	==> Switch systems
Main menu display options 5.4.1:	
[Exact approach main menu]	
Display general works	==> Options 5.4.2
Display narrower terms	==> Options 5.4.3
Display broader terms	==> Options 5.4.3
Display notes on this subject	==> Options 5.4.4
Display specialized subtopics	==> Options 5.4.5
Display specific places	==> Options 5.4.5
Display various time periods	==> Options 5.4.5
Display forms of publication	==> Options 5.4.5
Go to previous screen	==> Options 5.4.0
Expand search	==> Search trees
Start over	==> Switch systems
Bibliographic Record Display Options 5.4.2:	
[Bibliographic record display]	
Display next record	==> Options 5.4.2
Display previous record	==> Options 5.4.2
Display second page of record	==> Options 5.4.2
Exit record display	==> Options 5.4.1

Related Terms Display 5.4.3:	
[Broader or narrower terms display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected term	==> Options 5.4.1
Go back	==> Options 5.4.1
Scope Note Display Options 5.4.4:	
[Scope note display]	
Browse backward or forward	==> Manipulate scroll bar
Go back	==> Options 5.4.1
Subtopics Display Options 5.4.5:	
[Subtopics display]	
Browse backward or forward	==> Manipulate scroll bar
Search for selected subtopic	==> Options 5.4.1 or 5.4.5
Go back	==> Options 5.4.1

Figures 5.14–5.19 show the Blue System's response to the user query "trade and industry." In figure 5.14, the user enters the query "trade and industry" into the Blue System. The system's response to this query is an alphabetical keyword list of nine subject headings bearing the two words in the user query, i.e., "trade" and "industry" (figure 5.15).



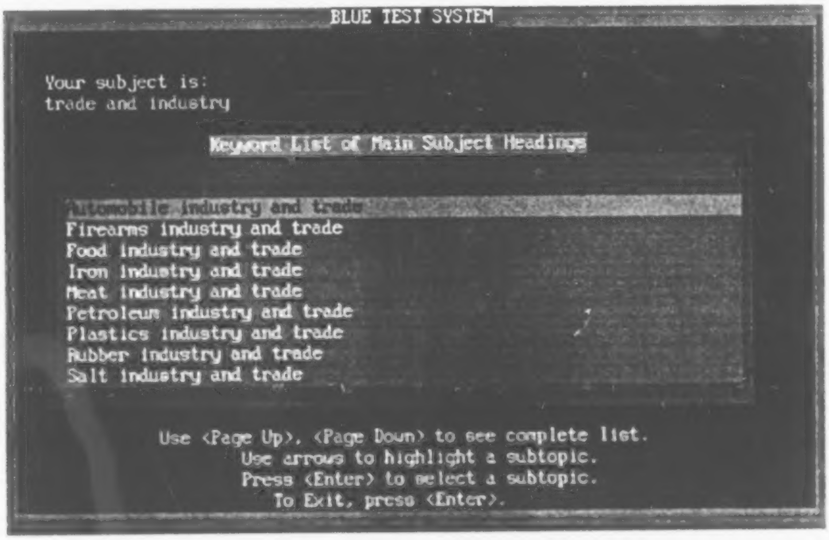
BLUE TEST SYSTEM

Please type in a subject to search and press <Enter>:

trade and industry

Figure 5.14. User entry of "control systems" query

BEST COPY AVAILABLE

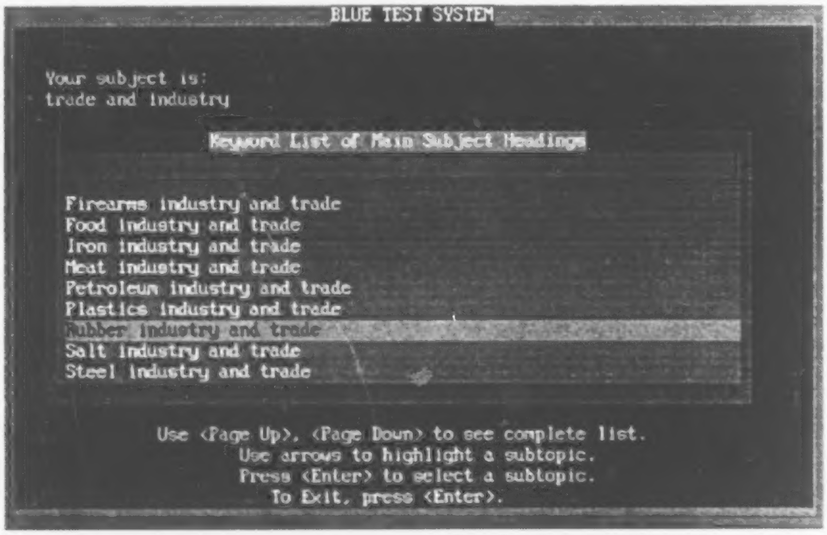


```
                    BLUE TEST SYSTEM

Your subject is:
trade and industry

          Keyword List of Main Subject Headings

  Automobile industry and trade
  Firearms industry and trade
  Food industry and trade
  Iron industry and trade
  Meat industry and trade
  Petroleum industry and trade
  Plastics industry and trade
  Rubber industry and trade
  Salt industry and trade

       Use <Page Up>, <Page Down> to see complete list.
            Use arrows to highlight a subtopic.
            Press <Enter> to select a subtopic.
               To Exit, press <Enter>.
```

Figure 5.15. Alphabetical keyword list for "trade and industry" query

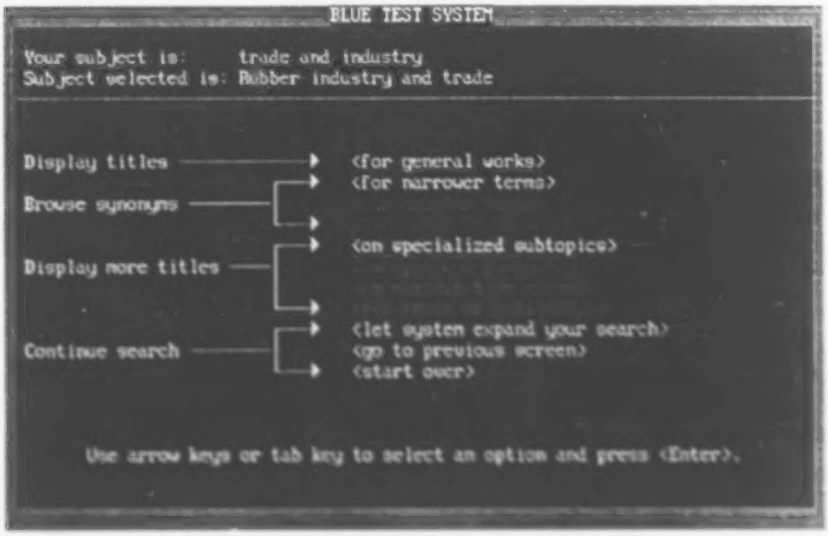
Since the alphabetical keyword list of subject headings only displays a maximum of nine subject headings at a time, the user has to press the <Page Down> key or click on the down arrow in the scroll bar to view the end of the list. In figure 5.16, the user selects the subject heading "Rubber industry and trade." In response, the Blue System conducts an exact search and produces the exact approach main menu bearing enabled options for "general works," "narrower terms," "specialized subtopics" (figure 5.17). If the user chooses the expand search option, the Blue System continues searching for the original query using the keyword-in-subdivided-heading approach which is the next approach in the search trees for queries exceeding one word. If the user chooses to "go to previous screen," the Blue System displays the original alphabetical keyword list (figure 5.15).

BEST COPY AVAILABLE



```
BLUE TEST SYSTEM

Your subject is:
trade and industry

Keyword List of Main Subject Headings

Firearms industry and trade
Food industry and trade
Iron industry and trade
Meat industry and trade
Petroleum industry and trade
Plastics industry and trade
Rubber industry and trade
Salt industry and trade
Steel industry and trade

Use <Page Up>, <Page Down> to see complete list.
Use arrows to highlight a subtopic.
Press <Enter> to select a subtopic.
To Exit, press <Enter>.
```

Figure 5.16. Selection of a listed heading from the alphabetical keyword list



```
BLUE TEST SYSTEM

Your subject is:     trade and industry
Subject selected is: Rubber industry and trade
```

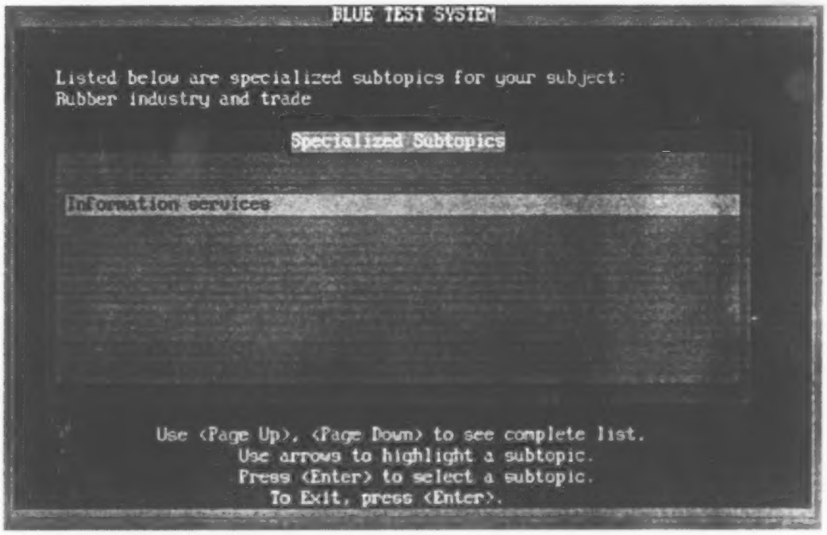
```
Use arrow keys or tab key to select an option and press <Enter>.
```

Figure 5.17. Exact approach main menu for "Rubber industry and trade"

Of the options listed in the exact approach main menu, the user selects the "specialized subtopics" option (figure 5.17). The Blue System responds with a single subtopic, "Information services," in figure 5.18. The user selects this subtopic and,

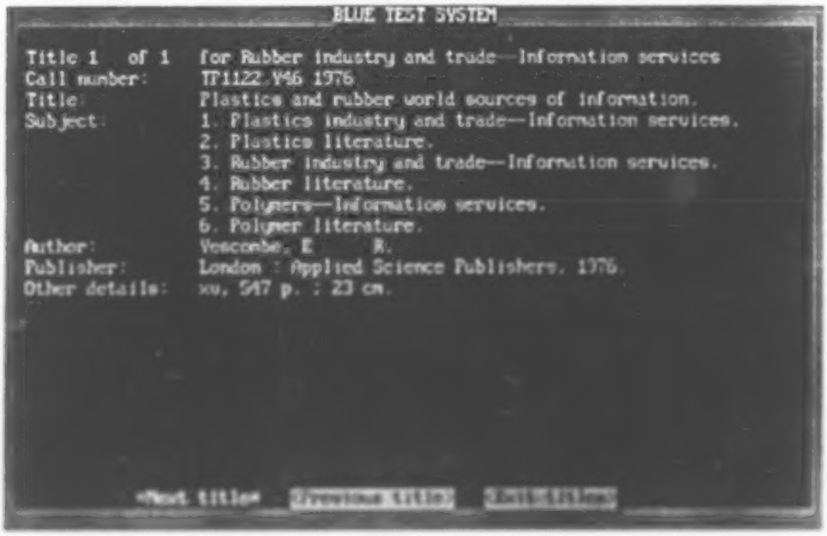
BEST COPY AVAILABLE

eventually, the Blue System displays the title shown in figure 5.19 that contains the subject heading "Rubber industry and trade — Information services."



```
                        BLUE TEST SYSTEM

    Listed below are specialized subtopics for your subject:
    Rubber industry and trade

                      Specialized Subtopics

     Information services

          Use <Page Up>, <Page Down> to see complete list.
                 Use arrows to highlight a subtopic.
                 Press <Enter> to select a subtopic.
                     To Exit, press <Enter>.
```

Figure 5.18. Selection of "Information services" subtopic in an exact search



```
                        BLUE TEST SYSTEM

  Title 1   of 1     for Rubber industry and trade—Information services
  Call number:       TP1122.V46 1976
  Title:             Plastics and rubber world sources of information.
  Subject:           1. Plastics industry and trade—Information services.
                     2. Plastics literature.
                     3. Rubber industry and trade—Information services.
                     4. Rubber literature.
                     5. Polymers—Information services.
                     6. Polymer literature.
  Author:            Vescombe, E     R.
  Publisher:         London : Applied Science Publishers, 1976.
  Other details:     xv, 547 p. : 23 cm.
```

Figure 5.19. Bibliographic record bearing subdivided heading "Rubber industry and trade"

BEST COPY AVAILABLE

5.6 Keyword-in-subdivided-heading Search

5.6.1 Background

The keyword-in-subdivided-heading search was implemented in both Blue and Pinstripe Test Systems. The two situations in which the Blue System invoked the keyword-in-subdivided-heading searches were similar to the situations in which this system invoked the keyword-in-main-heading search, viz. when user queries exceeded one word and failed to produce retrievals through the exact, alphabetical, and keyword-in-main-heading searches, and when user queries exceeded one word, produced retrievals through the exact, alphabetical, and/or keyword-in-main-heading searches, but users wished to continue searching. Neither situation guaranteed that keyword-in-subdivided-heading searches produced retrievals. The Pinstripe System randomly selected the keyword-in-subdivided-heading search through a random search selection algorithm.

When the keyword-in-subdivided-heading search yielded retrievals, the Blue and Pinstripe Systems responded with an alphabetical keyword list of subdivided subject headings bearing the words in the user query. The objective of displaying alphabetical keyword list of subject headings was the same as the objective of displaying an alphabetical list in the alphabetical approach and an alphabetical keyword list in the keyword-in-main-heading search, viz. to spark users' interest in a term that was partially but not entirely composed of words in the user query. When users selected a term from the alphabetical keyword list, the systems responded with a set of bibliographic records bearing the selected subject heading.

5.6.2 Generating Entries for the Keyword-in-subdivided-heading Search

The ASTUTE project team wrote a program to identify words in selected fields of bibliographic records and save them in a database supporting the various keyword approaches of the Blue and Pinstripe Systems. The program for the keyword-in-subdivided-heading search was the same as the program that handled words for the keyword-in-main-heading search (see section 5.5.2). The only difference was that the former program generated words for the database supporting the keyword-in-subdivided-heading search and the latter program generated words for the database supporting the keyword-in-main-heading search.

The keyword database supporting the keyword-in-subdivided-heading search contained words from the fields and subfields listed in Table 5.5.

Table 5.5. Fields/Subfields in the Keyword-in-subdivided-heading Search Database

USMARC Field Name	Format Type	Tag	Subfields
Subject (corporate body)	bibliographic	610	abndcxyz
Subject (meeting name)	bibliographic	611	andcktxyz
Subject (uniform title)	bibliographic	630	apsdxyz
Subject (topical)	bibliographic	650	axyz
Subject (geographic)	bibliographic	651	axyz

5.6.3 Responding to User Queries with the Keyword-in-subdivided-heading Search

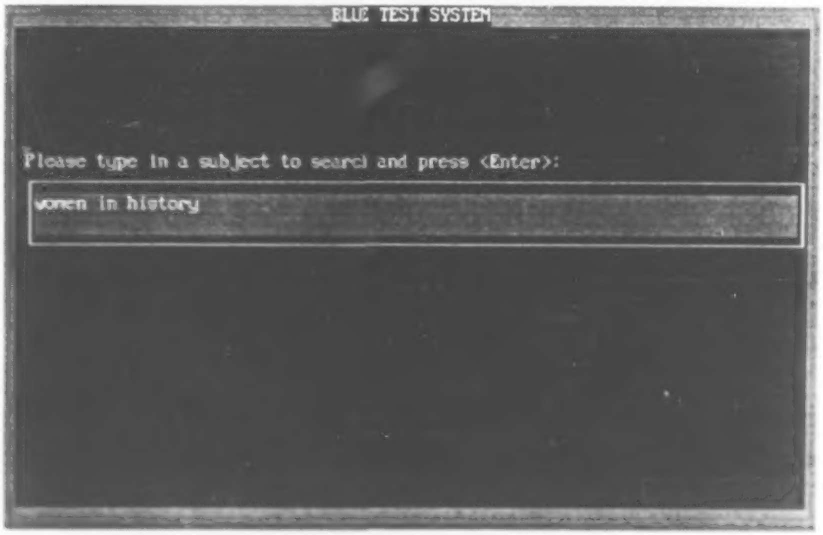
The keyword-in-subdivided-heading search was implemented in both Blue and Pinstripe Systems. With respect to the former, this search was the fourth search (after exact, alphabetical, and keyword-in-main-heading searches) that the Blue System enlisted to produce retrievals for user queries exceeding one word. When user queries matched words in main or subdivided subject headings, the Blue and Pinstripe Systems responded with an alphabetical keyword list of matched headings. User selection of a listed term resulted in a set of bibliographic records bearing the selected subject heading.

Table 5.6 lists options in two option boxes numbered 5.6.1 to 5.6.2. There are options for selecting a term from the alphabetical keyword list (options 5.6.1) and options for displaying bibliographic records (options 5.6.2).

Table 5.6. Keyword-in-subdivided-heading Search Approach Options

Alphabetical keyword list options 5.6.1:	
[Alphabetical term list]	
Choose listed heading	==> Options 5.6.2
Let system expand search	==> Search trees
Undo	==> Options 5.6.1
Start over	==> Switch systems

Bibliographic Record Display Options 5.6.2:	
[Bibliographic record display]	
Display next record	==> Options 5.6.2
Display previous record	==> Options 5.6.2
Display second page of record	==> Options 5.6.2
Exit record display	==> Options 5.6.1



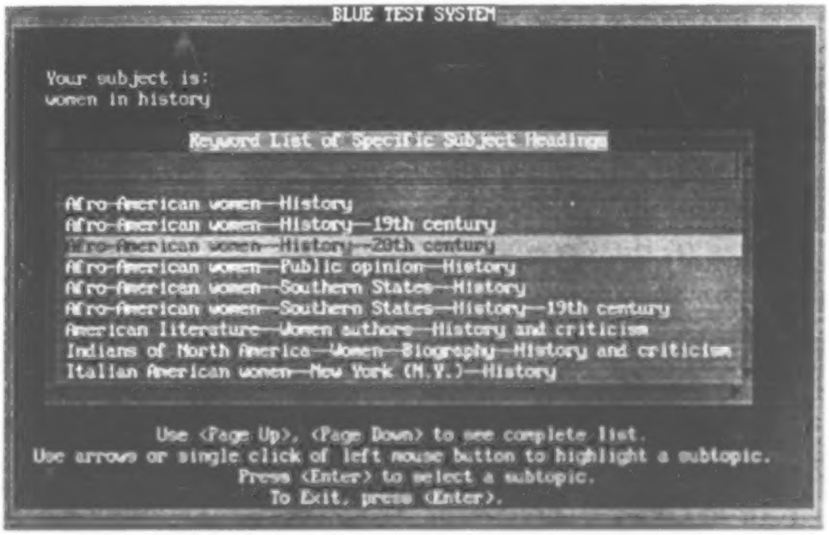
```
BLUE TEST SYSTEM

Please type in a subject to search and press <Enter>:
women in history
```

Figure 5.20. User entry of "women in history" query

Figures 5.20–5.23 show the Blue System's response to the user query "women in history." In figure 5.20, the user entersthe query "women in history" into the Blue System. The system's response to this query is an alphabetical keyword list of about two dozen subdivided subject headings bearing the two words in the user query, i.e., "women" and "history" (figure 5.21).

Since alphabetical keyword displays ofsubject headings are limited to a maximum of nine subject headings at a time, the usr has to press the «Page Down» key or click on the down arrow in the scroll bar to vie the end of the list (figure 5.21).

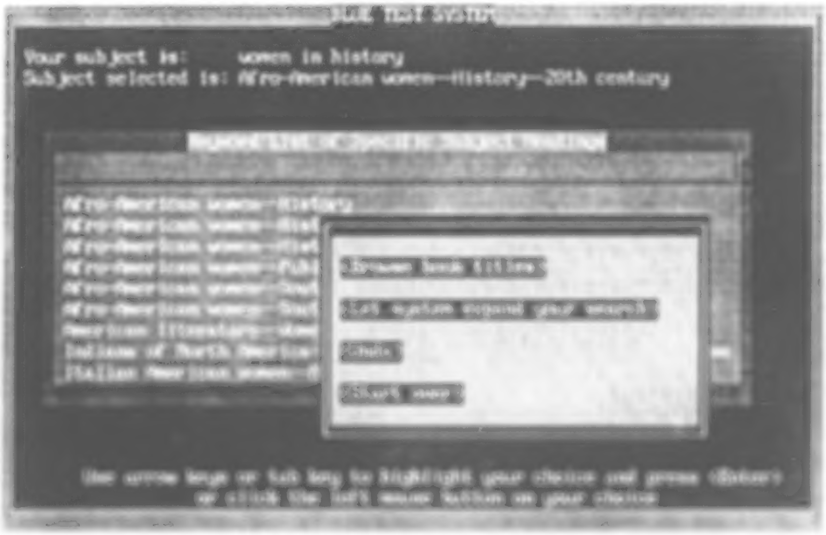
BEST COPY AVAILABLE



```
                         BLUE TEST SYSTEM

  Your subject is:
  women in history

            Keyword List of Specific Subject Headings

     Afro-American women--History
     Afro-American women--History--19th century
     Afro-American women--History--20th century
     Afro-American women--Public opinion--History
     Afro-American women--Southern States--History
     Afro-American women--Southern States--History--19th century
     American literature--Women authors--History and criticism
     Indians of North America--Women--Biography--History and criticism
     Italian American women--New York (N.Y.)--History

            Use <Page Up>, <Page Down> to see complete list.
  Use arrows or single click of left mouse button to highlight a subtopic.
                    Press <Enter> to select a subtopic.
                        To Exit, press <Enter>.
```

Figure 5.21. Alphabetical keyword list for "women in history" query

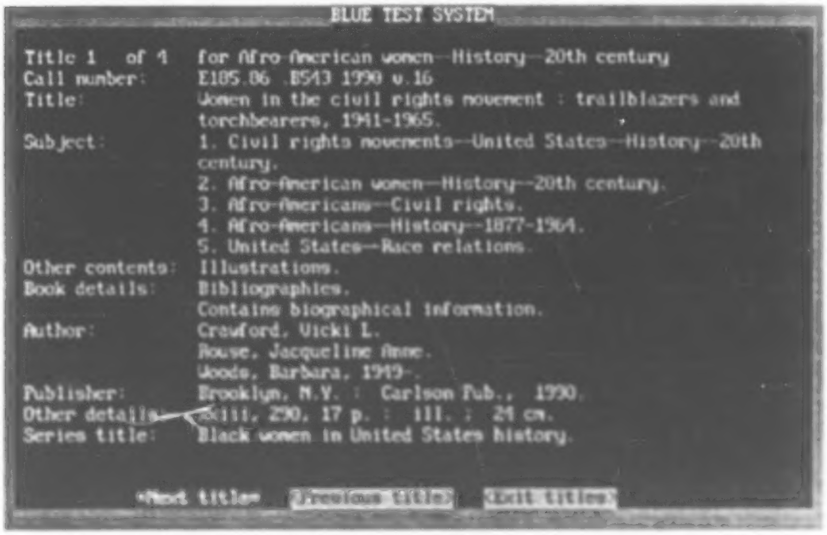
In figure 5.22, the user selects the subject heading "Afro-American women — History — 20th century." In response, the Blue System displays the first in a set of four titles (figure 5.23). The user can display other retrieved titles by clicking on "next title" and "previous title" options.



```
Your subject is:     women in history
Subject selected is: Afro-American women--History--20th century
```

Figure 5.22. User selection of subdivided heading

BEST COPY AVAILABLE



```
BLUE TEST SYSTEM

Title 1   of 4    for Afro-American women--History--20th century
Call number:      E185.86 .B543 1990 v.16
Title:            Women in the civil rights movement : trailblazers and
                  torchbearers, 1941-1965.
Subject:          1. Civil rights movements--United States--History--20th
                  century.
                  2. Afro-American women--History--20th century.
                  3. Afro-Americans--Civil rights.
                  4. Afro-Americans--History--1877-1964.
                  5. United States--Race relations.
Other contents:   Illustrations.
Book details:     Bibliographies.
                  Contains biographical information.
Author:           Crawford, Vicki L.
                  Rouse, Jacqueline Anne.
                  Woods, Barbara, 1949-.
Publisher:        Brooklyn, N.Y. :  Carlson Pub.,  1990.
Other details:    xxiii, 290, 17 p. :  ill. :  24 cm.
Series title:     Black women in United States history.
```

Figure 5.23. Bibliographic record bearing selected subdivided heading

To exit the title display, the user clicks on the "exit titles" option. This action returns the system to the alphabetical keyword list (figure 5.22).

5.7 Title-keyword Search

5.7.1 Background

The title-keyword search was only implemented in the Blue System. The search trees controlling this system called for the title-keyword approach in three situations: (1) when one-word user queries failed to produce retrievals through exact and alphabetical approaches, (2) when user queries exceeding one word failed to produce retrievals through the exact, alphabetical, and keyword-in-heading approaches, or (3) when user queries exceeded one word, produced retrievals through the exact, alphabetical, and/or keyword-in-heading approaches, but users wished to continue searching. None of these situations guaranteed that title-keyword searches would produce retrievals.

Prior to submitting queries exceeding one word to any keyword approach, the Blue System performed a keyword-in-record search of all subject-rich fields of bibliographic records for each word in the query. If it was unable to produce retrievals for one or more words of the user query, it displayed the query to users and prompted them to check their query's spelling. If the Blue System produced retrievals for every word of the user query, it ignored these intermediary results and returned control to the search tree for multi-word queries featuring keyword approaches (figure 4.5B).

5.7.2 Generating Entries for the Title-keyword Search

The ASTUTE project team wrote a program to identify words in selected fields of bibliographic records and save them in a database supporting the various keyword approaches of the Blue and Pinstripe Systems. The program for the title-keyword search was the same as the program that handled words for the keyword-in-main-heading search (see section 5.5.2). The only difference was that the former program generated words for the database supporting the title-keyword search and the latter program generated words for the database supporting the keyword-in-main-heading search.

The keyword database supporting the title-keyword search contained words from the fields and subfields listed in Table 5.7.

Table 5.7. Fields/Subfields in the Title-keyword Search Database

USMARC Field Name	Format Type	Tag	Subfields
Uniform title	bibliographic	130	afnpksd
Augmented title	bibliographic	214	a
Uniform title	bibliographic	240	anpks
Translation of title	bibliographic	242	anpb
Collective uniform title	bibliographic	243	a
Title statement	bibliographic	245	anpb
Subject (personal name)	bibliographic	600	tps
Subject (corporate name)	bibliographic	610	tps
Subject (meeting name)	bibliographic	611	andckt
Subject (uniform title)	bibliographic	630	tnk
Added entry (personal name)	bibliographic	700	ktnps
Added entry (corporate name)	bibliographic	710	tps
Added entry (meeting name)	bibliographic	711	andckt
Added entry (uniform title)	bibliographic	730	anpks
Added entry (title traced)	bibliographic	740	a
Variant (personal name)	bibliographic	870	t

5.7.3 Responding to User Queries with the Title-keyword Search

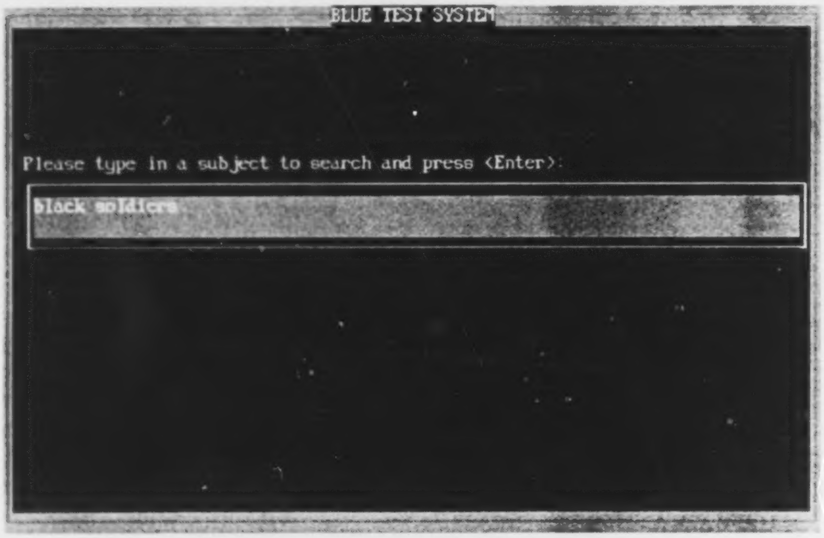
The title-keyword search was only implemented in Blue System. Title-keyword searches were the third and last approach given to one-word queries that failed to produce retrievals through exact and alphabetical approaches. Title-keyword searches were the fifth search (after exact, alphabetical, keyword-in-main-, and keyword-in-subdivided-heading searches) that the Blue System enlisted to produce retrievals for user queries exceeding one word. When user queries matched title words, the Blue System displayed the first in a set of bibliographic records bearing user query words.

A single option box is given table 5.8. The only options available in title-keyword searches were connected with displaying bibliographic records (options 5.8.1).

Table 5.8. Title-keyword Search Options

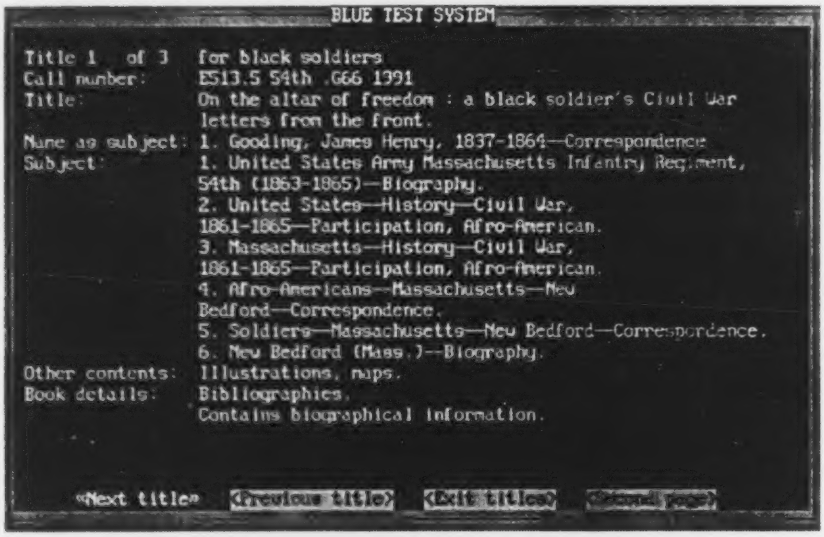
Bibliographic Record Display Options 5.8.1	
[Bibliographic record display]	
Display next record	==> Options 5.8.1
Display previous record	==> Options 5.8.1
Display second page of record	==> Options 5.8.1
Exit record display	==> Search trees

Figures 5.24–5.26 show the Blue System's response to the user query "black soldiers." In figure 5.24, the user enters the query "black soldiers" into the Blue System. Figure 5.25 shows the system's response to this query which is a display of the first of three retrieved titles.



```
BLUE TEST SYSTEM

Please type in a subject to search and press <Enter>:

black soldiers
```

Figure 5.24. User entry of "black soldiers" query

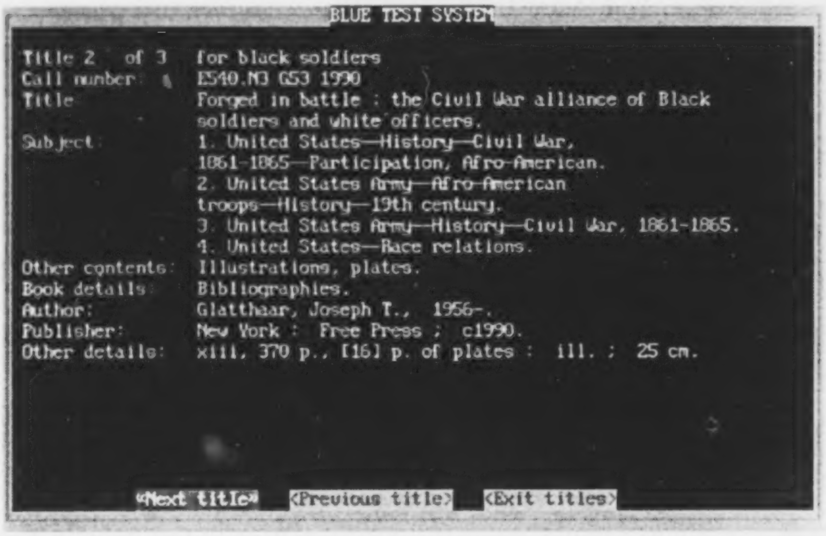


```
BLUE TEST SYSTEM
Title 1   of 3    for black soldiers
Call number:      E513.5 54th .G66 1991
Title:            On the altar of freedom : a black soldier's Civil War
                  letters from the front.
Name as subject:  1. Gooding, James Henry, 1837-1864--Correspondence
Subject:          1. United States Army Massachusetts Infantry Regiment,
                  54th (1863-1865)--Biography.
                  2. United States--History--Civil War,
                  1861-1865--Participation, Afro-American.
                  3. Massachusetts--History--Civil War,
                  1861-1865--Participation, Afro-American.
                  4. Afro-Americans--Massachusetts--New
                  Bedford--Correspondence.
                  5. Soldiers--Massachusetts--New Bedford--Correspondence.
                  6. New Bedford (Mass.)--Biography.
Other contents:   Illustrations, maps.
Book details:     Bibliographies.
                  Contains biographical information.
```

Figure 5.25. Bibliographic record bearing title words matching user query

Figure 5.26 shows the second of three retrieved titles. Note how both query words, i.e., "black," "soldiers," occur in the titles of the two bibliographic records.

BEST COPY AVAILABLE



```
Title 2   of 3   for black soldiers
Call number:     E540.N3 G53 1990
Title            Forged in battle : the Civil War alliance of Black
                 soldiers and white officers.
Subject:         1. United States—History—Civil War,
                 1861-1865—Participation, Afro-American.
                 2. United States Army—Afro-American
                 troops—History—19th century.
                 3. United States Army—History—Civil War, 1861-1865.
                 4. United States—Race relations.
Other contents:  Illustrations, plates.
Book details:    Bibliographies.
Author:          Glatthaar, Joseph T.,  1956-.
Publisher:       New York :  Free Press ;  c1990.
Other details:   xiii, 370 p., [16] p. of plates :  ill. ;  25 cm.
```

Figure 5.26. Bibliographic record bearing title words matching user query

To exit the title display, the user clicks on the "exit titles" option. Since the query exceeds one word, the Blue System gives the user the option to continue searching. If the user chooses to continue, the Blue System submits this query to a keyword search of subject heading fields and keyword-in-record search. If this query had been a one-word query, the Blue System would have switched to the Pinstripe System or initiated the post-search questionnaire.

5.8 Remaining Keyword Searches

5.8.1 Introduction

Two keyword searches remain: (1) keyword search of subject heading fields, and (2) keyword-in-record search. The former was unique to the Blue System. The latter was implemented in both Blue and Pinstripe Systems.

The two situations in which the Blue System invoked a keyword search of subject heading fields were similar to the situations in which this system invoked keyword-in-heading searches, that is, when user queries exceeded one word and failed to produce retrievals through previously executed searches, and when user queries exceeded one word and produced retrievals through previously executed searches but users wished to continue searching.

BEST COPY AVAILAB

When the Blue System was unable to produce retrievals through all previously executed searches (i.e., exact, alphabetical, keyword-in-heading, title-keyword, and keyword in subject heading fields), it submitted queries to keyword-in-record searches as a last resort. When users chose to continue searching after their queries produced retrievals through one or more previously executed searches, the Blue System submitted queries to keyword-in-record searches. The Pinstripe System randomly selected the keyword-in-record search through a random search selection algorithm.

5.8.2 Generating Entries for Remaining Keyword Searches

The ASTUTE project team wrote a program to identify words in selected fields of bibliographic records and save them in a database supporting the various keyword approaches of the Blue and Pinstripe Systems. The programs for keyword in subject heading field and keyword-in-record searches were the same as the program that handled the keyword-in-main-heading search (see section 5.5.2). The only difference was that the former programs generated words for two separate databases that supported the keyword search of subject heading fields and keyword-in-record searches, respectively.

The keyword databases supporting the keyword-in-record search contained words from the fields and subfields listed in Table 5.9. Fields marked with an asterisk also contributed to the keyword database supporting the keyword search of subject heading fields.

Table 5.9. Fields/Subfields in the Keyword-in-record Database

USMARC Field Name	Format Type	Tag	Subfields
Illustration codes	bibliographic	008	positions 18–21
Nature of contents codes	bibliographic	008	positions 24–27
Biography code	bibliographic	008	position 34
Festschrift code	bibliographic	008	position 30
Main entry (corporate name)	bibliographic	110	abncktp
Main entry (meeting name)	bibliographic	111	andc
Main entry (uniform title)	bibliographic	130	atnpksd
Augmented title	bibliographic	214	a
Uniform title	bibliographic	240	anpks
Translation of title	bibliographic	242	anpb
Collective uniform title	bibliographic	243	a

Title statement	bibliographic	245	anpb
Series statement added entry	bibliographic	440	ap
Series statement (untraced)	bibliographic	490	a
Contents note	bibliographic	505	a
Summary	bibliographic	520	a
Subject (personal name)	bibliographic	600	tps
*Subject (corporate name)	bibliographic	610	abndcxyztps
*Subject (meeting names)	bibliographic	611	andcktxyz
*Subject (uniform title)	bibliographic	630	apsdxyztnk
*Subject (topical)	bibliographic	650	axyz
*Subject (geographic)	bibliographic	651	axyz
Subject (uncontrolled)	bibliographic	653	a
Subject (genre/form)	bibliographic	655	axyz
Added entry (personal name)	bibliographic	700	ktnps
Added entry (corporate name)	bibliographic	710	tps
Added entry (meeting name)	bibliographic	711	ancktp
Added entry (uniform title)	bibliographic	730	atnpks
Added entry (title traced)	bibliographic	740	a
Added entry (physical characteristics)	bibliographic	755	axyz
Series (personal name)	bibliographic	800	a
Series (corporate name)	bibliographic	810	t
Series (meeting name)	bibliographic	811	at
Series (uniform title)	bibliographic	830	at
Variant (personal name)	bibliographic	870	t
Variant (corporate name)	bibliographic	871	t
Variant (meeting name)	bibliographic	872	t
Variant (uniform title)	bibliographic	873	a

*Fields/subfields contributing to keyword search of subject heading fields database.

5.8.3 Responding to User Queries with Remaining Keyword Searches

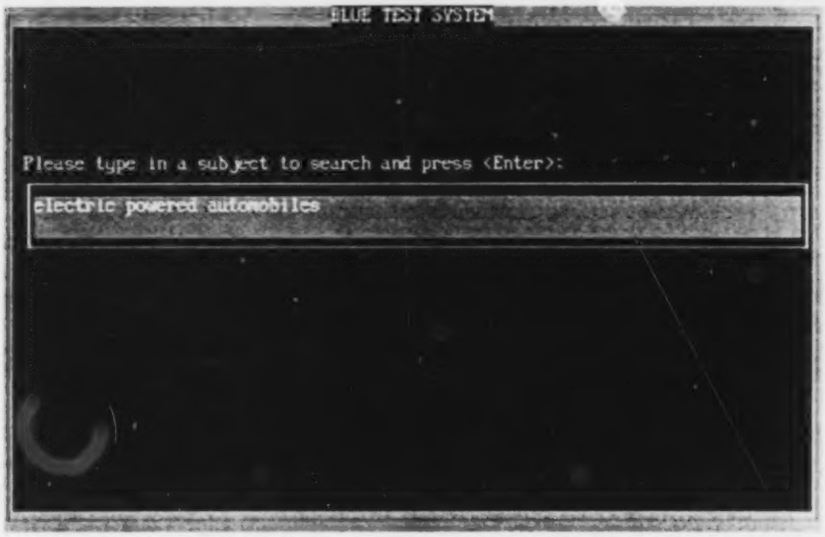
The keyword search of subject heading fields was only implemented in Blue System. User queries exceeding one word that failed to produce retrievals through five other searches were given to this search. Keyword-in-record searches were the search of last resort for queries exceeding one word. When keyword-in-record searches failed to make retrievals in the Blue System, the system notified users of its failure. The system then switched to the Pinstripe System (if users had not begun in the Pinstripe System) or initiated the post-search questionnaire (if users had begun in the Pinstripe System). When user queries produced retrievals through the keyword search of subject heading fields (Blue System only) or keyword-in-record searches (Blue and Pinstripe Systems), the system displayed the first in a set of bibliographic records bearing user query words.

Table 5.10 enumerates options available to searchers in keyword searches of subject heading fields or keyword-in-record searches. The only options available in these searches are connected with displaying bibliographic records (options 5.10.1).

Table 5.10. Subject Heading Keyword and Keyword-in-record Search Options

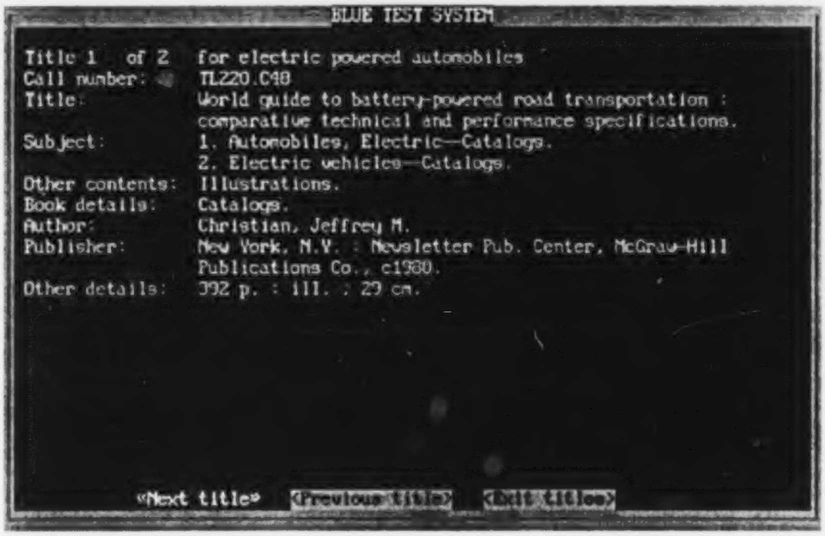
Bibliographic Record Display Options 5.10.1	
[Bibliographic record display]	
Display next record	==> Options 5.10.1
Display previous record	==> Options 5.10.1
Display second page of record	==> Options 5.10.1
Exit record display	==> Search trees

Figures 5.27–5.28 show the Blue System's response to the user query "electric powered automobiles." In figure 5.27, the user enters the query "electric powered automobiles" into the Blue System. Figure 5.28 shows the system's response to this query which is a display of the first of two retrieved titles. In the first record, user query words match a word in the title, i.e., "powered," and words in subject headings, i.e., "electric" and "automobiles."



```
Please type in a subject to search and press <Enter>:

electric powered automobiles
```

Figure 5.27. User entry of "electric powered automobiles" query



```
Title 1   of 2   for electric powered automobiles
Call number:     TL220.C48
Title:           World guide to battery-powered road transportation :
                 comparative technical and performance specifications.
Subject:         1. Automobiles, Electric—Catalogs.
                 2. Electric vehicles—Catalogs.
Other contents:  Illustrations.
Book details:    Catalogs.
Author:          Christian, Jeffrey M.
Publisher:       New York, N.Y. : Newsletter Pub. Center, McGraw-Hill
                 Publications Co., c1980.
Other details:   392 p. : ill. ; 29 cm.
```

Figure 5.28. Bibliographic record bearing words matching user query

BEST COPY AVAILABLE

5.9 Personal-name Searches

5.9.1 Background

The functionality of personal-name searches must take into account the content and structure of assigned subject headings for personal names. Personal-name headings come from the assigned subject headings in libraries' bibliographic records. Catalogers establish these headings by following AACR2 rules and guidelines and previous LC practice. They refer to the Library of Congress Name Authority File (LCNAF) to verify personal names. References are generated from LCNAF records for names used in libraries' bibliographic records. Catalogers also refer to SCM:SH to add subdivisions to personal-name headings. The results are single assigned subject headings bearing elements for personal names, topical subjects, and, possibly, geographic names. The order of elements in such headings are: (1) last name, (2) first name or initial, (3) possibly, a middle name or initial, (4) possibly, one or more dates, (5) possibly, topical subdivisions, and (6) rarely, geographic subdivisions. Subdivisions appended to personal-name headings are usually for topical subjects; however, several subdivisions under names of persons authorize use of geographic subdivisions (Library of Congress 1990, H1110).

The empirical study of user queries which served as the impetus for the formulation of search trees tested in this research project demonstrated that users rarely expressed all six elements of personal-name headings in their queries (Drabenstott and Vizine-Goetz, 1994). Furthermore, when two or more elements were included, users rarely entered elements in the sequence prescribed by personal-name headings.

To improve the performance of personal-name subject searching, systems must ask users whether their queries involve personal names. When users identify such queries, systems also need users to provide them with information about the particular elements included in queries to distinguish personal-name elements from topical-subject elements. Systems can then accommodate the rigid structure of personal-name headings by searching for as many elements as needed to provide useful retrievals.

Recommendations from the empirical study regarding subject searches for personal-name subject queries involved three subject searching approaches, i.e., keyword-in-subdivided heading, keyword-in-record, and alphabetical searches. The types of searches systems undertake depend upon the elements users offer in their subject queries for personal names. Systems present results for only those searches that produce postings in the library catalog. Table 4.1 lists the sequence of the three recommended subject searching approaches and the query elements each approach enlists to produce retrievals. If an approach fails to produce retrievals, systems proceed with the next approach and element(s) on the list.

At the most, systems need three elements, viz. last names, first names, and topics, to perform keyword searches of single subdivided headings. If such searches fail to produce retrievals, systems drop first names and perform keyword-in-subdivided-heading searches for the remaining two elements. Keyword searches of all subject-rich fields of bibliographic records may yield retrievals for queries that are combinations of personal names and topics. If keyword approaches fail to produce retrievals, the system omits the topical element(s) and performs an alphabetical search using as many name elements as match assigned subject headings and references in the catalog.

For example, the three elements "kennedy" (last name), "robert" (first name), and "presidential primaries" (topic), would probably not be successful retrieving single assigned subject headings bearing these words, i.e., first two listed approaches in Table 4.1, because the topic elements are not used in subdivisions under subject headings for personal names. Keyword-in-record approaches may yield bibliographic records bearing the topic elements and name elements because titles and/or note fields may contain the topic elements, viz. "presidential primaries." Retrieved records may have assigned subject headings such as "Kennedy, Robert Francis, 1925–1968 — Career in politics" or "Kennedy, Robert Francis, 1925–1968 — Political and social views." If users are not satisfied with the results, systems could continue with the alphabetical approach but they should omit the topical element of the query to effect a match of a personal-name subject heading.

When users do not enter a topical element(s), the alphabetical approach is recommended. Systems respond with an alphabetical list of assigned subject headings for personal names in the alphabetical neighborhood of the entered last name. Alphabetical lists could be limited to the name subfields of personal-name subject headings, i.e., subfields $a, $q, $b, $c, and $d. Subject headings subdivided by topical or geographic subdivisions would not be included to reduce the amount of information on screens. Alphabetical indexes of personal-name subject headings that are limited to name subfields are sufficient for letting users find and verify the name they are looking for. Once they select a main heading, systems can then display the topical and geographic subdivisions appended to it through the exact approach or display the bibliographic records bearing unsubdivided and subdivided forms of the selected heading.

When users enter a last name only, systems also respond with the alphabetical approach. Alphabetical lists are also the last resort when systems are unable to find matches of two or more query elements. For example, a query with two elements "nixon" (last name) and "prosecution" (topic) may not retrieve single headings or bibliographic records, but an alphabetical search for the last name "nixon" will probably place the user at a location in the index where she can browse forward to find the assigned subject heading "Nixon, Richard Milhous, 1913– 1994." Selecting this heading will give her the opportunity to display assigned headings with subdivisions such as "Impeachment" or "Pardon" which may be appropriate to this topic.

The alphabetical approach recommended here is similar to approaches described by other researchers who have studied personal-name queries (Dickson 1984, 36; Taylor 1984, 15). They recommended that systems first find assigned headings matching surnames and given names in user queries. If they fail, systems should then look for assigned headings matching surnames and the first letter of given names. Lastly, systems should look for assigned headings that match the surnames users enter.

On their own, online systems cannot distinguish between query terms that are topical subjects and terms that are personal names. Systems need to enlist users' assistance to identify the various elements of personal-name queries. They can then make useful responses to user queries even though the structure of controlled vocabulary terms and user queries differ.

5.9.2 Keyword Searches for Personal Names

The ASTUTE project team limited keyword searching in the Blue and Pinstripe Test Systems to a single keyword search, i.e., keyword-in-record searches, for two reasons. First, UM-Dearborn bibliographic records contained few (about a dozen) subdivided subject headings for personal names. Thus, the likelihood that a user query would produce retrievals as a result of keyword-in subdivided-heading searches was so low that the team felt such searches would be a waste time and computer resources. Second, most subdivided subject headings for personal names in Earlham College bibliographic records were unique, i.e., assigned to a single bibliographic record. Thus, a successful keyword-in-subdivided-heading search would produce an intermediary display of the single retrieved heading which would eventually result in the display of a single bibliographic record. Since keyword-in-record searches would retrieve this same record, and, possibly, additional one(s), the project team decided to streamline keyword searches for personal names by implementing a single search — the keyword-in-record search. Keyword-in-subdivided-heading searches of personal name and topic elements would be appropriate for much larger databases.

The ASTUTE project team used the same program to identify words in selected fields of bibliographic records and save them in a database supporting the keyword-in-record search for personal names as the program had written for other keyword approaches (see section 5.5.2). The keyword database supporting this search contains words from the fields and subfields listed in Table 5.11.

Table 5.11. Fields/Subfields in the Personal-name, Keyword-in-Record Database

USMARC Field Name	Format Type	Tag	Subfields
Illustration codes	bibliographic	008	positions 18–21
Nature of contents codes	bibliographic	008	positions 24–27
Biography code	bibliographic	008	position 34
Festschrift code	bibliographic	008	position 30
Main entry (corporate name)	bibliographic	110	abncktp
Main entry (meeting name)	bibliographic	111	andc
Main entry (uniform title)	bibliographic	130	atnpksd
Augmented title	bibliographic	214	a
Uniform title	bibliographic	240	anpks
Translation of title	bibliographic	242	anpb
Collective uniform title	bibliographic	243	a
Title statement	bibliographic	245	anpb
Series statement added entry	bibliographic	440	ap
Series statement (untraced)	bibliographic	490	a
Contents note	bibliographic	505	a
Summary	bibliographic	520	a
Subject (personal name)	bibliographic	600	aqbcktxyznps
Subject (corporate name)	bibliographic	610	abndcxyztps
Subject (meeting names)	bibliographic	611	andcktxyz
Subject (uniform title)	bibliographic	630	apsdxyztnk
Subject (topical)	bibliographic	650	axyz
Subject (geographic)	bibliographic	651	axyz
Subject (uncontrolled)	bibliographic	653	a
Subject (genre/form)	bibliographic	655	axyz
Added entry (personal name)	bibliographic	700	ktnps
Added entry (corporate name)	bibliographic	710	tps
Added entry (meeting name)	bibliographic	711	ancktp
Added entry (uniform title)	bibliographic	730	atnpks
Added entry (title traced)	bibliographic	740	a
Added entry (physical characteristics)	bibliographic	755	axyz

Series (personal name)	bibliographic	800	a
Series (corporate name)	bibliographic	810	t
Series (meeting name)	bibliographic	811	at
Series (uniform title)	bibliographic	830	at
Variant (personal name)	bibliographic	870	t
Variant (corporate name)	bibliographic	871	t
Variant (meeting name)	bibliographic	872	t
Variant (uniform title)	bibliographic	873	a

Keyword-in-record searches for personal-name queries were implemented in both Blue and Pinstripe Systems. The Blue System performed a keyword-in-record search using first name, last name, and topic elements. If it failed to yield retrievals, it omitted the first name element and tried again using only last name and topic elements.

A random selection algorithm governed the Pinstripe System's selection of the keyword-in-record search. The Pinstripe System performed a keyword-in-record search using all user-query elements; it did not omit query elements to effect retrievals.

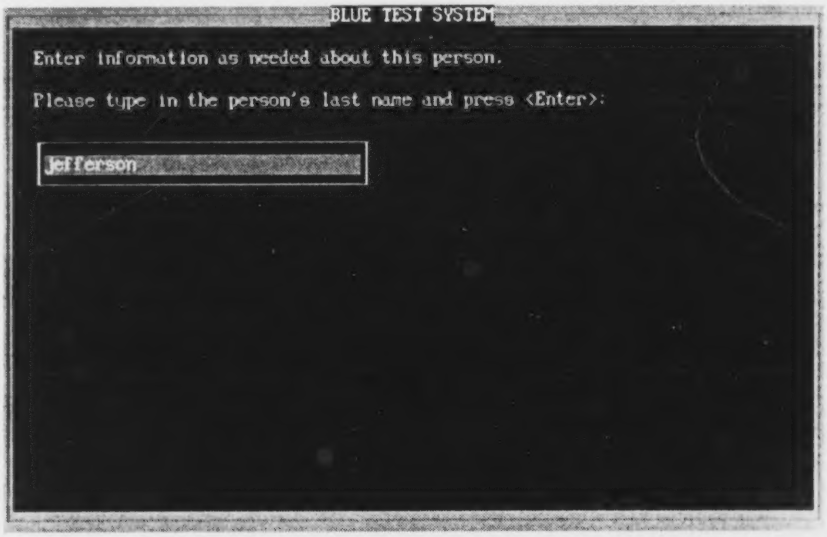
The single option box is given in table 5.12 for keyword-in-record searches. When this search produced retrievals, the system displayed the first record and provided options to display the next record, previous record, second page of a record, or exited the record display.

Table 5.12. Personal-name, Keyword-in-record Search Options

Bibliographic Record Display Options 5.12.1	
[Bibliographic record display]	
Display next record	==> Options 5.12.1
Display previous record	==> Options 5.12.1
Display second page of record	==> Options 5.12.1
Exit record display	==> Search trees

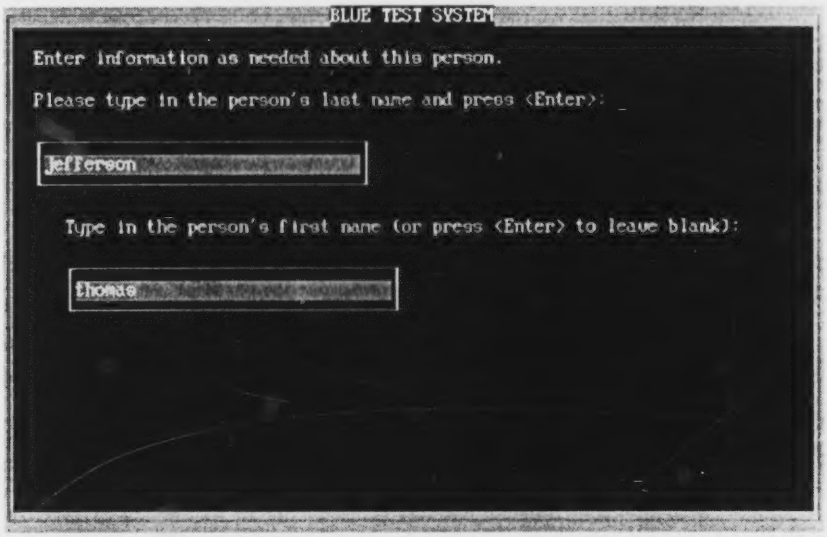
Figures 5.29–5.32 show the Blue System's response to the user query bearing elements for "jefferson" (last name), "thomas" (first name), and "on slavery." In figure 5.29, the user enters the last name element of the query "jefferson" into the Blue System. The user's pressing of the <Enter> key prompts the Blue System to respond with a second pop-up box requesting the first name element of the query. In figure 5.30, the user enters the first name element "thomas." When the user presses the <Enter> key, the

system prompts the user to enter a topic element. In figure 5.31, the user enters the topic element of the query.



```
BLUE TEST SYSTEM
Enter information as needed about this person.
Please type in the person's last name and press <Enter>:
Jefferson
```

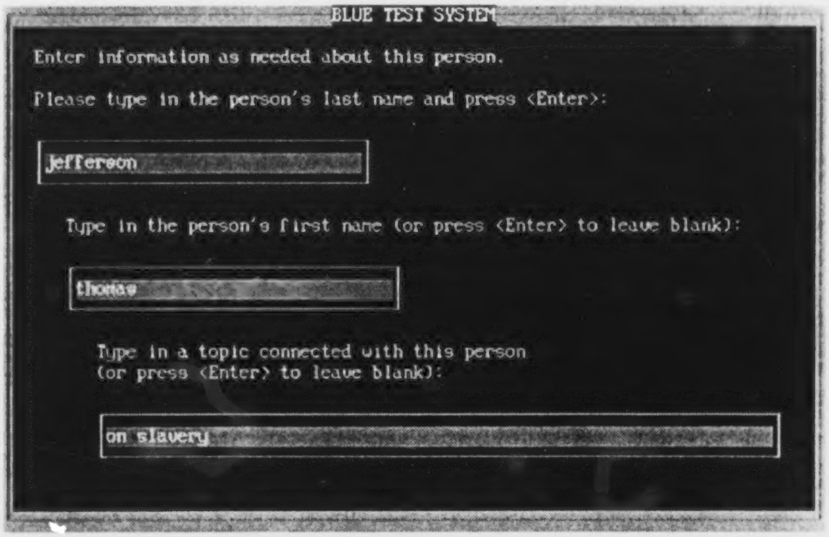
Figure 5.29. User entry of last name query element



```
BLUE TEST SYSTEM
Enter information as needed about this person.
Please type in the person's last name and press <Enter>: _
Jefferson
  Type in the person's first name (or press <Enter> to leave blank):
  thomas
```

Figure 5.30. User entry of first name query element

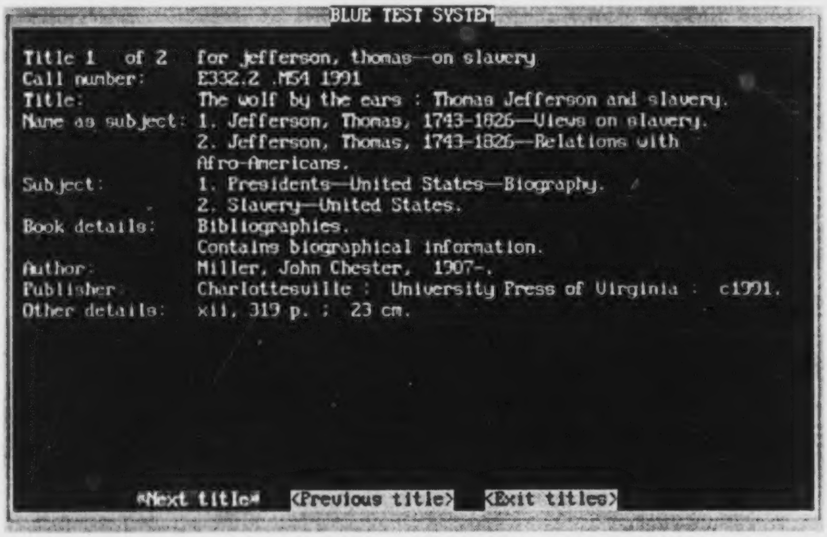
BEST COPY AVAILABLE



```
Enter information as needed about this person.

Please type in the person's last name and press <Enter>:

jefferson

  Type in the person's first name (or press <Enter> to leave blank):

  thomas

    Type in a topic connected with this person
    (or press <Enter> to leave blank):

    on slavery
```

Figure 5.31. User entry of topic query element

The Blue System retrieves two titles bearing the words in the user query. Figure 5.32 shows the first of the two retrieved titles.



```
Title 1  of 2   for jefferson, thomas--on slavery
Call number:    E332.2 .M54 1991
Title:          The wolf by the ears : Thomas Jefferson and slavery.
Name as subject: 1. Jefferson, Thomas, 1743-1826--Views on slavery.
                2. Jefferson, Thomas, 1743-1826--Relations with
                Afro-Americans.
Subject:        1. Presidents--United States--Biography.
                2. Slavery--United States.
Book details:   Bibliographies.
                Contains biographical information.
Author:         Miller, John Chester,  1907-.
Publisher:      Charlottesville :  University Press of Virginia :  c1991.
Other details:  xii, 319 p. ;  23 cm.
```

Figure 5.32. Bibliographic record bearing words in user query

When the Blue-system user is done reviewing retrieved bibliographic records, the system prompts the user to continue searching. The system responds to users who

BEST COPY AVAILABLE

choose to continue searching with the results of an alphabetical search. If the user continues the search shown in figures 5.29–5.32, the Blue System would display an alphabetical list of personal names. Choosing the name heading "Jefferson, Thomas, 1743–1826" results in 62 titles. The user would have to review the records to determine whether certain titles might discuss Jefferson's views on slavery.

The Pinstripe System does not give users an opportunity to continue searching. It switches to the Blue System (if the Pinstripe System is the first system in which the user conducts a search) or initiates the post-search questionnaire (if the user has already searched the Blue System for the particular query).

5.9.3 Alphabetical Approach for Personal Names

The alphabetical approach was only implemented in the Blue System. The Blue System performed an alphabetical search in three situations: (1) when keyword-in-record searches failed to produce retrievals, (2) when keyword-in-record searches produced retrievals but the user chose to continue searching, and (3) when users only entered name elements in their personal-name subject queries.

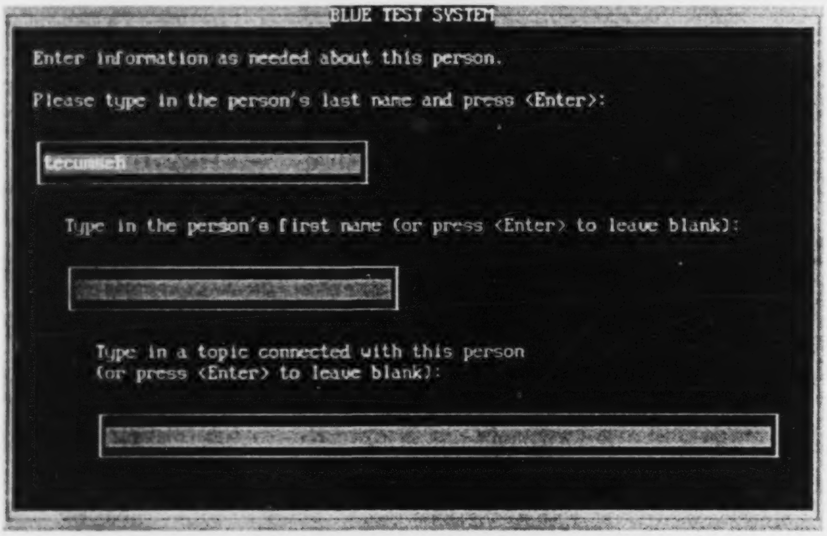
Only the personal name subject heading field (USMARC tag 600) contributed to the alphabetical list of name headings. Terms in subfields $a, $b, $q, $c, and $d were displayed to users.

Table 5.13 enumerates options available to searchers in alphabetical searches of subject headings for personal names. Options were available for manipulating the alphabetical subject list (options 5.13.0) and displaying retrieved bibliographic records (options 5.13.1).

Table 5.13. Alphabetical Search Approach Options for Personal-name Queries

Alphabetical name selection options 5.13.0:	
[Alphabetical name list]	
Choose listed name heading	==> Options 5.13.1
Browse backward or forward in list	==> Click on scroll bar
Undo	==> Options 5.13.0
Start over	==> Switch systems
Bibliographic Record Display Options 5.13.1:	
[Bibliographic record display]	
Display next record	==> Options 5.13.1
Display previous record	==> Options 5.13.1
Display second page of record	==> Options 5.13.1
Exit record display	==> Switch systems

Figure 5.33–5.36 show the Blue System's response to the user query "tecumseh." In figure 5.33, the user enters this personal-name subject query. The Blue System responds with an alphabetical list bearing the personal-name subject heading "Tecumseh, Shawnee Chief, 1763–1813" (figure 5.34).



BLUE TEST SYSTEM

Enter information as needed about this person.

Please type in the person's last name and press <Enter>:

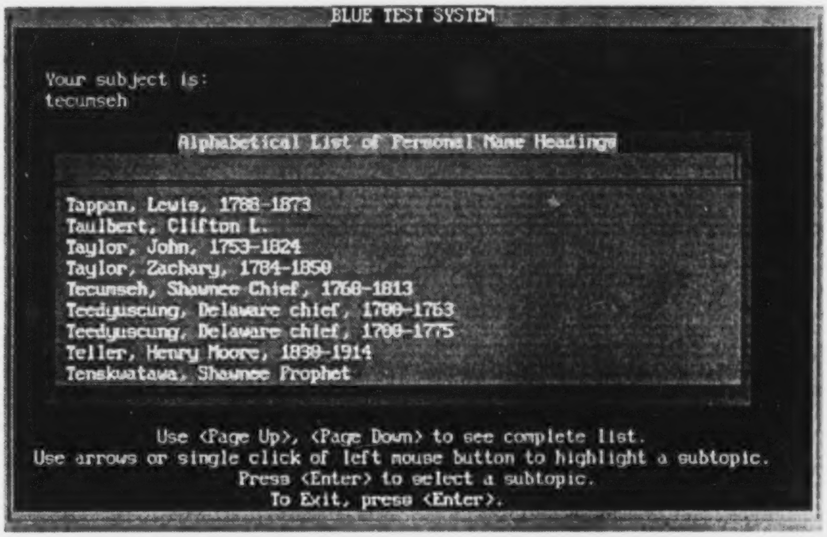
tecumseh

Type in the person's first name (or press <Enter> to leave blank):

Type in a topic connected with this person
(or press <Enter> to leave blank):

Figure 5.33. User entry of personal-name query for "tecumseh"

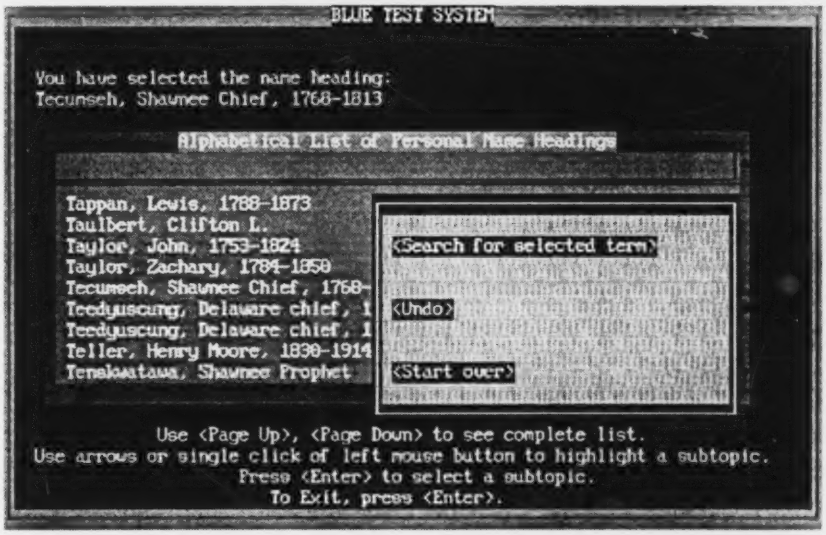
BEST COPY AVAILABLE



```
                          BLUE TEST SYSTEM

  Your subject is:
  tecumseh

              Alphabetical List of Personal Name Headings

    Tappan, Lewis, 1788-1873
    Taulbert, Clifton L.
    Taylor, John, 1753-1824
    Taylor, Zachary, 1784-1850
    Tecumseh, Shawnee Chief, 1768-1813
    Teedyuscung, Delaware chief, 1700-1763
    Teedyuscung, Delaware chief, 1700-1775
    Teller, Henry Moore, 1830-1914
    Tenskwatawa, Shawnee Prophet

            Use <Page Up>, <Page Down> to see complete list.
 Use arrows or single click of left mouse button to highlight a subtopic.
                     Press <Enter> to select a subtopic.
                        To Exit, press <Enter>.
```

Figure 5.34. Alphabetical list of personal-name subject headings in response to "tecumseh" query

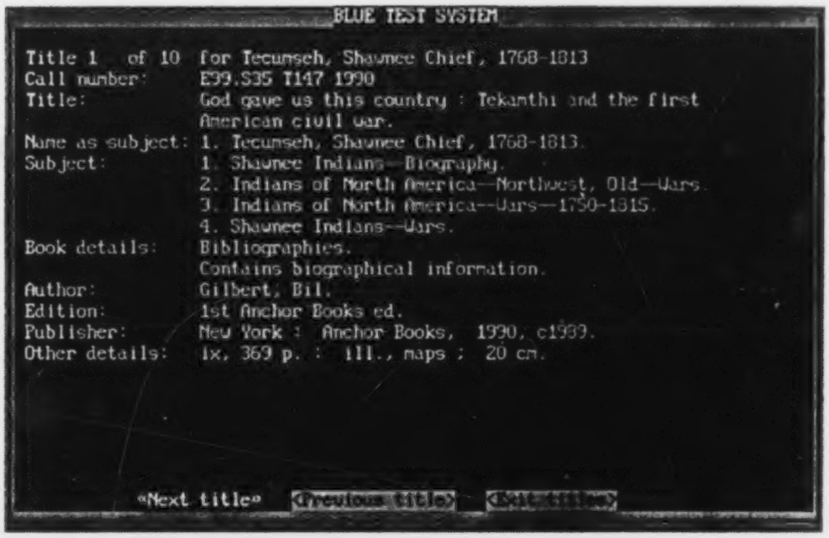
Figure 5.35 shows the user selecting the Tecumseh subject heading from the alphabetical list of personal-name subject headings. The Blue System's response to the user's selection is to display the first of ten retrieved titles (figure 5.36).



```
                          BLUE TEST SYSTEM

  You have selected the name heading:
  Tecumseh, Shawnee Chief, 1768-1813

              Alphabetical List of Personal Name Headings

    Tappan, Lewis, 1788-1873
    Taulbert, Clifton L.
    Taylor, John, 1753-1824
    Taylor, Zachary, 1784-1850
    Tecumseh, Shawnee Chief, 1768-
    Teedyuscung, Delaware chief, 1
    Teedyuscung, Delaware chief, 1
    Teller, Henry Moore, 1830-1914
    Tenskwatawa, Shawnee Prophet

            Use <Page Up>, <Page Down> to see complete list.
 Use arrows or single click of left mouse button to highlight a subtopic.
                     Press <Enter> to select a subtopic.
                        To Exit, press <Enter>.
```

Figure 5.35. User selection of a personal-name subject heading

BEST COPY AVAILABLE



```
Title 1   of 10  for Tecumseh, Shawnee Chief, 1768-1813
Call number:     E99.S35 T147 1990
Title:           God gave us this country : Tekamthi and the first
                 American civil war.
Name as subject: 1. Tecumseh, Shawnee Chief, 1768-1813.
Subject:         1. Shawnee Indians--Biography.
                 2. Indians of North America--Northwest, Old--Wars.
                 3. Indians of North America--Wars--1750-1815.
                 4. Shawnee Indians--Wars.
Book details:    Bibliographies.
                 Contains biographical information.
Author:          Gilbert, Bil.
Edition:         1st Anchor Books ed.
Publisher:       New York :  Anchor Books,  1990, c1989.
Other details:   ix, 369 p. :  ill., maps ;  20 cm.
```

Figure 5.36. Bibliographic record bearing the subject heading "Tecumseh"

5.10 Pinstripe System's Random Selection Algorithm

The Pinstripe System featured three subject searching approaches for subjects generally: (1) alphabetical, (2) keyword-in-subdivided-heading, and (3) keyword-in-record searches. The implementation of these three searches was almost the same as the implementation of these searches in the Blue System. The only difference was that the Pinstripe System's selection of a subject searching approach in response to a user query was governed by a random selection algorithm.

The Pinstripe System's random selection algorithm used a random function provided by FoxPro to facilitate the selection of an approach. The alphabetical approach was assigned a value range from 1 to 100, the keyword-in-subdivided-heading search was assigned a range from 101 to 200, and the keyword-in-record approach was assigned a range from 201 to 300.

The first step in selecting a search was for the Pinstripe System to call the random function which returned a decimal number between 0 and 1. This number was multiplied by 300 and the product was compared to the range for each approach to select a search type. Let's say that the random function called the decimal number 0.28. The system figured the product of 0.28 by 300 equaled 84. It compared 84 to the range for the three searches and selected the alphabetical search because it fell into the 1 to 100 range.

BEST COPY AVAILABLE

The Pinstripe System featured only the keyword-in-record search for subject searches for personal names. This system did not delete topical elements from personal-name queries to effect matches.

5.11 Displaying Bibliographic Records

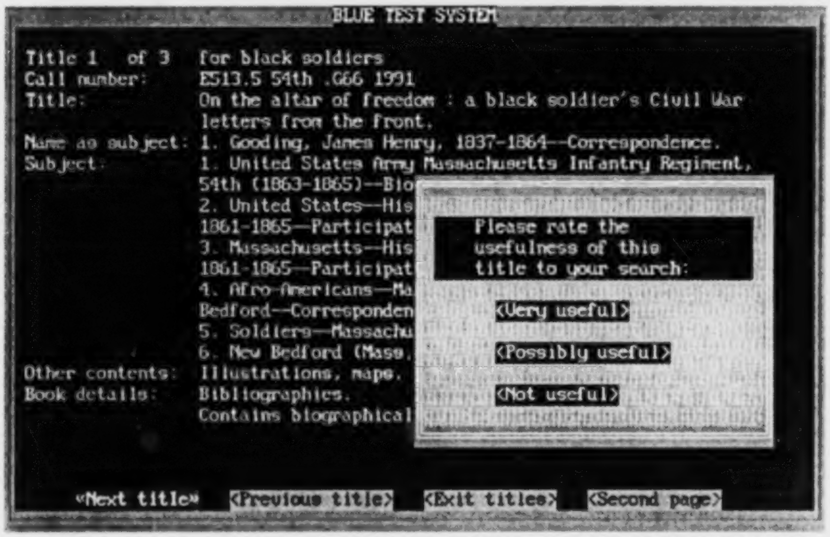
The Blue and Pinstripe Systems displayed bibliographic records and collected relevance assessments from end users in the same way. Table 5.14 lists fields and subfields that contributed text to bibliographic record displays. Names listed in the "USMARC Field Name" column of Table 5.14 were same names as the field labels that introduced such data in bibliographic record displays. They were also listed in the same order as they appeared in such displays.

Table 5.14. Fields/Subfields for Bibliographic Record Displays

USMARC Field Name	Format Type	Tag	Subfields
Call number	bibliographic	099	ab
		092	ab
		050	ab
Title	bibliographic	245	apb
Alternate title	bibliographic	130	atnpksd
		214	a
		240	anpks
		242	anpb
		243	a
		730	atnpks
		740	a
		873	a
Name as subject	bibliographic	600	aqbcdktxyznps
Subject	bibliographic	610	abndcxy
		611	andcktxyz
		630	apsdxyztnk
		650	axyz
		651	axyz
Summary	bibliographic	505	a
Contents	bibliographic	520	a
Other contents	bibliographic	008	positions 18–21 (replace codes with text equivalent)

Book details	bibliographic	870	t
		871	t
		008	positions 24-27, 30, 34 (replace codes with text equivalent)
Author	bibliographic	100	aqbcde
		110	abndcktp
		700	aqbcdektnps
		710	abncdtps
Sponsor	bibliographic	111	andc
		711	andcktp
Edition	bibliographic	250	a
Publisher	bibliographic	260	abc
Other details	bibliographic	300	abc
Series title	bibliographic	440	ap
Other series	bibliographic	490	a
		800	a
		810	t
		811	at
		830	at

Figures 5.5, 5.13, 5.19, 5.23, 5.25, 5.26, 5.28, 5.32, and 5.36 showed bibliographic record displays in the ASTUTE experimental online catalog. The system gave users three options following record displays: (1) display next title, (2) display previous title, or (3) exit title display. When users chose any of these options, the system produced a pop-up window that asked users to rate the usefulness of the displayed title. For example, let's say that the user who displayed the title in figure 5.25 selected the "next title" option. Before displaying the next title, the system would produce the pop-up box shown in figure 5.37.



```
                              BLUE TEST SYSTEM
Title 1   of 3    for black soldiers
Call number:      E513.5 54th .G66 1991
Title:            On the altar of freedom : a black soldier's Civil War
                  letters from the front.
Name as subject:  1. Gooding, James Henry, 1837-1864--Correspondence.
Subject:          1. United States Army Massachusetts Infantry Regiment,
                  54th (1863-1865)--Bio
                  2. United States--His
                  1861-1865--Participat
                  3. Massachusetts--His
                  1861-1865--Participat
                  4. Afro-Americans--Ma
                  Bedford--Corresponden
                  5. Soldiers--Massachu
                  6. New Bedford (Mass.
Other contents:   Illustrations, maps.
Book details:     Bibliographies.
                  Contains biographical

                  Please rate the
                  usefulness of this
                  title to your search:
```

Figure 5.37. Prompting users for relevance assessments

The pop-up window in figure 5.37 prompts users to rate the title "very useful," "somewhat useful," or "not useful."

5.12 Chapter Summary

This chapter provided a non-technical discussion of ASTUTE functionality. ASTUTE was composed of two experimental online catalogs: (1) the Blue System in which search trees controlled the system's selection of a subject searching approach in response to user queries, and (2) the Pinstripe System which chose subject searching approaches randomly in response to user queries. It gave the background and rationale for subject searching approaches in the experimental systems, described the fields that contributed words and phrases to searchable FoxPro databases, discussed the process of the various subject searches, and gave examples of subject searches in the Blue and Pinstripe Systems to demonstrate functionality and system responses to user queries.

References

Chan, Lois Mai. 1986. *Library of Congress Subject Headings: principles and application.* Littleton, CO: Libraries Unlimited.

Cochrane, Pauline A. 1986. *Improving LCSH for use in online catalogs.* Littleton, CO: Libraries Unlimited.

BEST COPY AVAILABLE

Conway, Martha O'Hara, ed. 1992. *The future of subdivisions in the Library of Congress Subject Headings system.* Washington, DC: Library of Congress.

Dickson, Jean. 1984. "An analysis of user errors in searching an online catalog." *Cataloging & Classification Quarterly* 4, 3 (Spring): 19–38.

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Holley, Robert P., and Robert E. Killheffer. 1982. "Is there an answer to the subject access crisis?" *Cataloging & Classification Quarterly* 1(2/3): 125–33.

Library of Congress. 1990. *Subject cataloging manual: subject headings.* Washington, DC: The Library.

Markey, Karen. 1984. *Subject searching in library catalogs: Before and after the introduction of online catalogs.* Dublin, Ohio: OCLC.

Massicotte, Mia. 1988. "Improved browsable displays for online subject access." *Information Technology and Libraries* 7, 4 (December): 373–80.

Taylor, Arlene G. 1984. "Authority files in online catalogs: An investigation of their value." *Cataloging & Classification Quarterly* 4, 3 (Spring): 1–19.

Van Pulis, Noelle, and Lorene E. Ludy. 1988. "Subject searching in an online catalog with authority control." *College & Research Libraries* 49 (6): 523–33.

Wiberley, Stephen E., and Robert Allen Daugherty. 1988. "Users' persistence in scanning lists of references." *College & Research Libraries* 49 (March): 149–56.

6 Methods Used to Test the New Subject Access Design

6.1 Introduction

The ASTUTE experimental online catalog was actually composed of two online catalogs: (1) the Blue Test System, in which search trees governed the system's selection of a subject searching capability, and (2) the Pinstripe Test System, in which the system selected a subject searching capability randomly. These systems were purposely designed to be very much alike to focus the attention of library patrons and staff on the retrieval of useful information in response to their queries. The Blue and Pinstripe Systems had virtually the same interfaces, and they accessed the same bibliographic and authority databases. Except for the Blue System's enhancement with the search trees, the two systems and their capabilities were the same.

In online retrieval tests, the Blue and Pinstripe Systems encouraged library patrons and staff to perform online subject searches in both systems. In introductory screens, the systems informed users that they were participating in an experiment and allowed them to graciously exit the system if they did not want to take part. Prior to conducting subject searches, the systems asked participants to respond to three pre-search questions that collected demographic information. During their searches, the systems recorded user activity and user relevance assessments for displayed bibliographic records to a time-stamped transaction log. The systems asked participants to answer eleven questions in a post-search questionnaire that compared the performance of the Blue and Pinstripe Test Systems. Thus, the retrieval tests yielded quantitative comparative data about the systems' capabilities and effectiveness, and about users' system preference. After completing the post-search questionnaire, library staff were asked additional questions that prompted them to make comparisons between the two systems with respect to the systems' ease of use, subject searching capabilities, and performance. These qualitative data were combined with quantitative data from patron searches to increase our understanding of search tree performance and enhancement.

In this chapter, the experimental design of the Comparison Search Experiment is described. Data collection procedures, online questionnaires (administered to patrons and staff), and offline questionnaires (administered to staff) are highlighted.

6.2 Previous Studies Comparing Systems

The comparative method used in this project is based on the Comparison Search Experiment devised by Siegel et al. (1983) in their evaluation of two prototype online catalogs at the National Library of Medicine (NLM). The goal of Siegel's Comparison Search Experiment was to reproduce realistic search conditions in which library patrons performed a search on a topic of their own choosing on both systems. The researchers alternated the first system patrons used to minimize any transfer effect, that is, bias caused by patrons transferring their knowledge of the search and of searching in the first system to the second system. The researchers observed patrons and decided when they should switch systems. After completing their search of the second system, the researchers asked patrons questions about their experiences using the two systems, satisfaction with search results, and system preference. Siegel et al. used a second approach — the Sample Search Experiment — to control for potentially confounding variables. Siegel et al. used the results of both experiments to select an online catalog for use by library patrons and reference staff at NLM.

Several researchers have based their evaluation of two or more online catalogs on Siegel's Comparison or Sample Search Experiment. The project director used Siegel's methods to evaluate two experimental online catalogs, one with enhanced subject searching functionality using the Dewey Decimal Classification (DDC) and one without the enhancement (Markey and Demeyer 1986, 109–18). She expanded on Siegel's approach by asking patrons to assess the relevance of retrieved bibliographic records and used these assessments to compare the performance of the two experimental online catalogs with respect to precision and estimated recall. Markey and Demeyer used the results of their comparative study to evaluate the retrieval effectiveness of the DDC in subject searches.

Jones (1988) enlisted the Comparison Search Experiment to compare the performance of the operational LIBERTAS online catalog and the experimental Okapi online catalog. In the two separate studies by Markey and Demeyer (1986) and Jones (1988), interviewers approached prospective online catalog users, asked them to participate in the study, asked them a few questions prior to their search, observed their online searches in two systems, and asked them much longer lists of questions after the search about their searching experiences, satisfaction with search results, and system preference. Jones collected user relevance assessments and used the data to establish user assessments of system performance and user attitudes to Okapi's recall-improvement devices. Jones also compared user preferences between specific interaction features connected with ease of use, bibliographic format, browsing, menu design, and learning and relearning how to use the system.

Since its inception in 1982, the Okapi experimental online catalog has been under continual enhancement and redesign. Enhancement and redesign has been based on the findings of many evaluations including Jones' 1988 evaluation that enlisted the Comparison Search Experiment.

In 1987, the Okapi designers evaluated three different Okapi representations that were alike in all ways except for the following functionality: (1) the experimental (EXP) system featured "strong" stemming, a phrase dictionary, and some automatic cross-referencing, (2) the control (CTL) system featured "weak" stemming and spelling standardization, and (3) the third system (OSTEM) performed no stemming or spelling normalization (Walker and Jones 1987). The EXP and CTL systems were available for patron searching on different floors of the library. Although interviewers were present when patrons searched the EXP or CTL systems, their interaction with patrons was minimal. They recorded start and finish times and asked patrons four brief questions about themselves and their search results. Okapi logged patron searches to transaction logs. The researchers performed a failure analysis of EXP and CTL searches using questionnaire responses and logged searches. They also searched OSTEM for interesting searches and compared OSTEM results with patron searches in EXP or CTL. The goal of the Okapi designers in this particular evaluation was to assess the performance of several recall-improvement devices, i.e., automatic word stemming, synonym and cross-reference tables, soundex keys for matching personal names, and an n-gram technique for approximate word matching.

Based on 1987 and 1988 evaluations, the Okapi designers set to work at new representations of this experimental system. In 1990, they evaluated three different Okapi representations: (1) the DUMB system that was the same as the EXP system in the 1987 evaluation, (2) the QE (i.e., query expansion) system that was the same as "dumb" except for a "look for books similar to those already chosen" feature, and (3) the FULL system that was the same as QE except for a shelf-browsing feature (Walker and De Vere 1990). The Okapi researchers devised a method that "lay somewhere between the Comparison and Sample search experiments used in the NLM experiment" (Walker and De Vere 1990, 46). They recruited respondents, asked them a few questions about themselves before their searches, gave them a choice of questions to search on two of the three systems, recorded searches to transaction logs, asked them more questions about their searching experiences, results, and system preference after the search, and obtained relevance assessments for displayed bibliographic records from judges such as librarians and subject specialists. The goal of the evaluation was to compare the three systems with regard to effectiveness, efficiency, and user acceptability.

Hildreth (1993) employed an approach similar to the 1990 Okapi evaluation to evaluate the usability and retrieval performance of a navigation approach to subject searching in online catalogs. He evaluated three different experimental online catalogs: (1) a control catalog bearing no navigation features, (2) a catalog that supported related-record navigation based on title words, and (3) a catalog that supported related-record navigation based on subject headings only. He recruited respondents, asked them a few questions about themselves before their search, gave them a search task to perform in the three systems, recorded searches to transaction logs, asked them more questions about their searching experiences, results, and system preference after the search, and compared their search results with those obtained by expert searchers.

6.3 Online Questionnaire Administration in Catalog Use Studies

Comparative approaches to system evaluation have used human intermediaries to collect information about searchers and their searching experiences. Such approaches are expensive and time-consuming for several reasons. Human intermediaries are needed to recruit searchers for participation in the system evaluation. An interviewer must be present to monitor end-user searches, ask questions, and record answers. If searchers are not recruited and scheduled ahead of time, interviewers could spend a considerable amount of time recruiting searchers themselves and explaining the study to them. Travel expenses for interviewers can be considerable for extended data-collection periods at off-campus sites. For these reasons, the project director used a comparative approach to system evaluation in this project that did not use interviewers or human intermediaries at all. The experimental online catalog performed and/or recorded all activity connected with data collection — searcher recruitment, questionnaire administration, search logging, and logging of searcher relevance assessments.

The comparative approach used in this study was innovative and experimental. Although online questionnaire administration was used in an early online catalog use study, delegating complete control of retrieval tests to the experimental online catalog had not been done before.

The early study of online catalog use that enlisted online questionnaire administration was conducted by researchers at the University of California's Division of Library

Automation (UC/DLA) who administered online questionnaires to users of the MELVYL online catalog in the Council on Library Resources-sponsored nationwide study of online catalogs (University of California 1982). Despite the success of UC/DLA's online questionnaire administration, this approach has been rarely used in online catalog use studies.

In a recent Department of Education-sponsored study, a research team at Rutgers University recorded searches and online questionnaire responses administered in online pre- and post-search questionnaires to transaction logs (Belkin et al. 1990). The purpose of the Rutgers study was to discover the goals, tasks, and behaviors of library and online catalog users with a view to using these data for the specification of online catalog design principles.

A research team at City University collected data on online catalog use using methods compatible with those enlisted by the Rutgers team (Hancock-Beaulieu, McKenzie, and Irving 1990). The City University team developed Olive, a front-end microcomputer program that recorded searches and pre-, in-, and post-search online questionnaire responses. Olive also featured search playback and printing capabilities. Although the City University team recorded searches and questionnaire responses made by users of two different online catalogs (i.e., CLSI and LIBERTAS), the team's aim was to build on previous research findings regarding predictors and categories of user behavior.

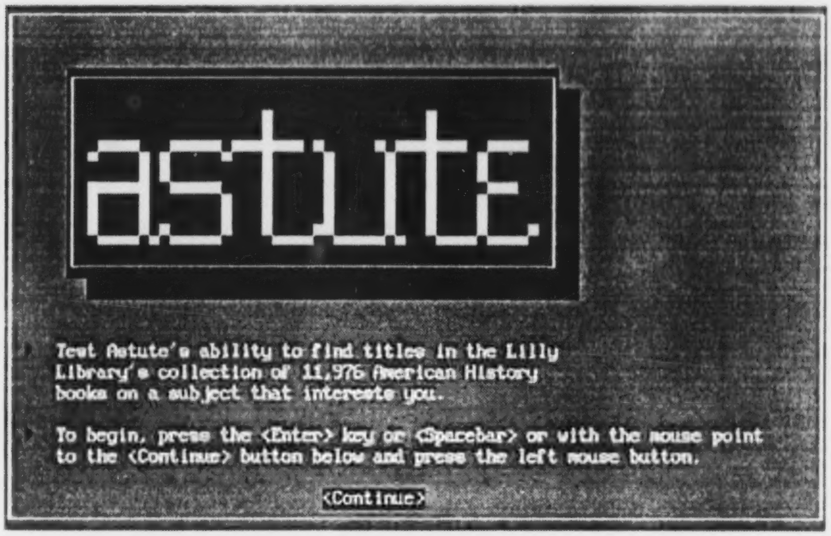
6.4 Comparison Search Experiment (Library Patrons)

6.4.1 Recruiting Searchers

In this study, the researchers transported and assembled a Gateway microcomputer bearing ASTUTE to the two data collection sites — Mardigian Library at the University of Michigan-Dearborn and Lilly Library at Earlham College. The microcomputer was dedicated to use of the ASTUTE experimental online catalog. At UM-Dearborn, ASTUTE was located in a quiet study area of the library which was also near computer science, engineering, and technology stacks. Thus, ASTUTE searchers would not have to go very far to access the library material they retrieved in their searches of the experimental online catalog. At Earlham College, ASTUTE was located in the reference area of the library near the library's MARCIVE™ CD/ROM-based online catalog and other CD/ROM reference sources. Lilly Library reference staff were also nearby and directed patrons to ASTUTE when they felt patrons would find useful material in the system. At both libraries, signs were placed near ASTUTE to attract library patrons to use the system.

The ASTUTE experimental online catalog performed recruiting functions on its own. The presence of the microcomputer equipment in a public place in the library, posted written signs, and two alternating introductory screen savers attracted potential respondents. Reading subsequent introductory screens also piqued the interest or curiosity of some individuals to take part in the system evaluation. ASTUTE was installed in Mardigian and Lilly Libraries for several weeks at a time. If interested or curious library users did not have time to use the system the first time they saw it, they might have used it on subsequent library visits when they had more time.

ASTUTE featured five introductory screens. The first two introductory screens also functioned as screen savers to avoid the image of one of the screens burning into the monitor. Figure 6.1 shows the first introductory screen that invited users to take part in the evaluation of the system.



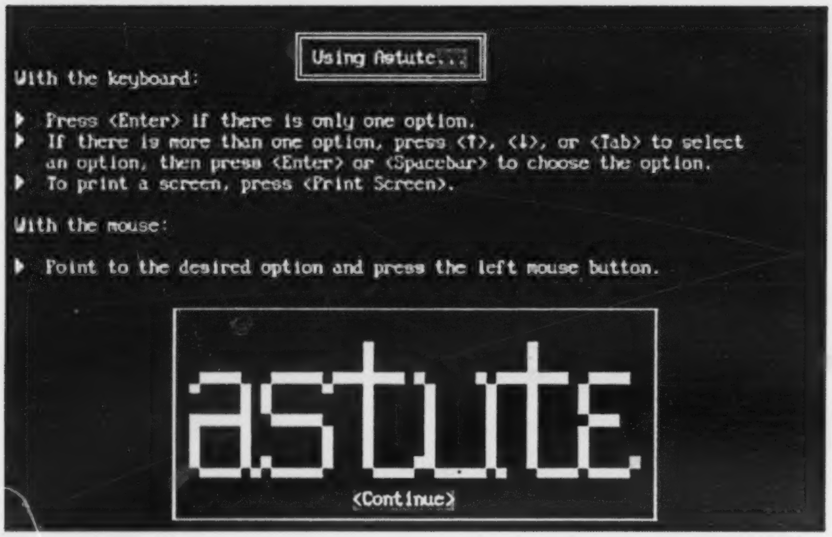
Test Astute's ability to find titles in the Lilly Library's collection of 11,976 American History books on a subject that interests you.

To begin, press the <Enter> key or <Spacebar> or with the mouse point to the <Continue> button below and press the left mouse button.

Figure 6.1 First introductory screen

Figure 6.2 shows the second introductory screen that told users how to operate the keyboard and mouse, make selections, and print screens.

BEST COPY AVAILABLE



With the keyboard:

- Press <Enter> if there is only one option.
- If there is more than one option, press <↑>, <↓>, or <Tab> to select an option, then press <Enter> or <Spacebar> to choose the option.
- To print a screen, press <Print Screen>.

With the mouse:

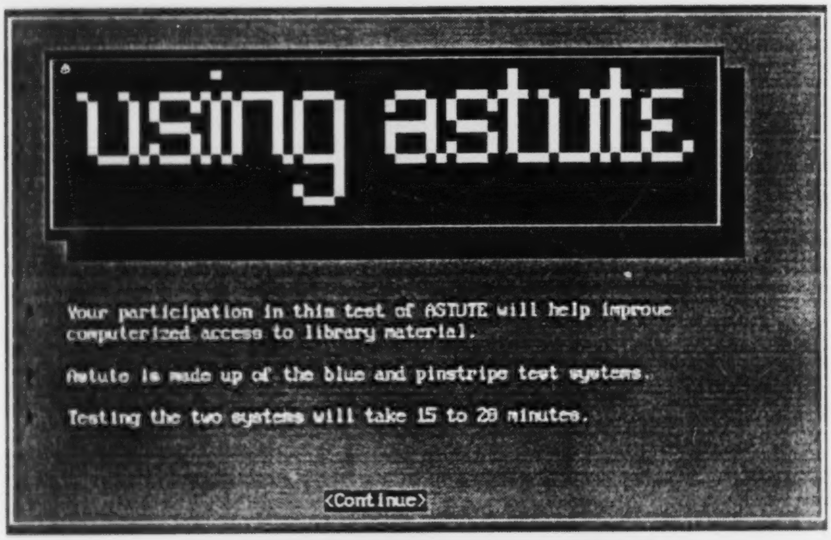
- Point to the desired option and press the left mouse button.

Figure 6.2 Second introductory screen

If no one was using ASTUTE, the system alternated between displaying the first two introductory screens in figures 6.1. and 6.2.

When users clicked on the <Continue> button in the first two introductory screens, ASTUTE responded with a series of three additional introductory screens that explained their participation in the test of the system. Prospective respondents were encouraged to read these three screens to find out about the system's purpose, subject coverage, and capabilities. These screens also invited users to conduct an online search on a topic of their own choosing in the system. ASTUTE told users it was logging their searches and responses to questions. Library users were entirely on their own to read screens, conduct searches, and answer questions. Figures 6.3–6.5 show the series of three additional introductory screens.

BEST COPY AVAILABLE

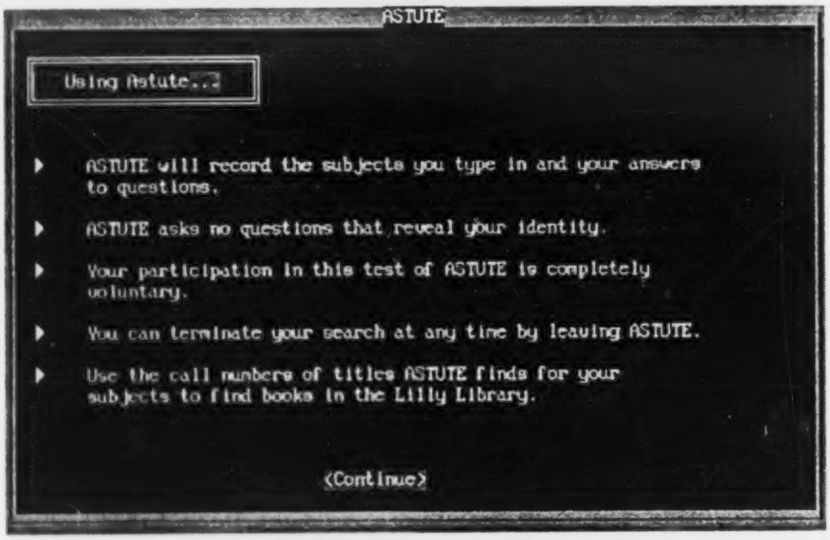


Your participation in this test of ASTUTE will help improve computerized access to library material.

Astute is made up of the blue and pinstripe test systems.

Testing the two systems will take 15 to 20 minutes.

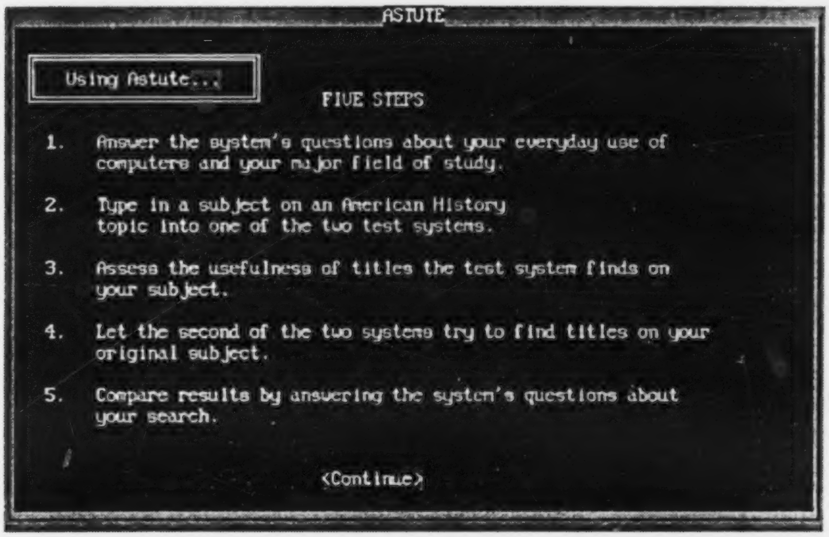
Figure 6.3. Third introductory screen



- ASTUTE will record the subjects you type in and your answers to questions.
- ASTUTE asks no questions that reveal your identity.
- Your participation in this test of ASTUTE is completely voluntary.
- You can terminate your search at any time by leaving ASTUTE.
- Use the call numbers of titles ASTUTE finds for your subjects to find books in the Lilly Library.

Figure 6.4. Fourth introductory screen

BEST COPY AVAILABLE



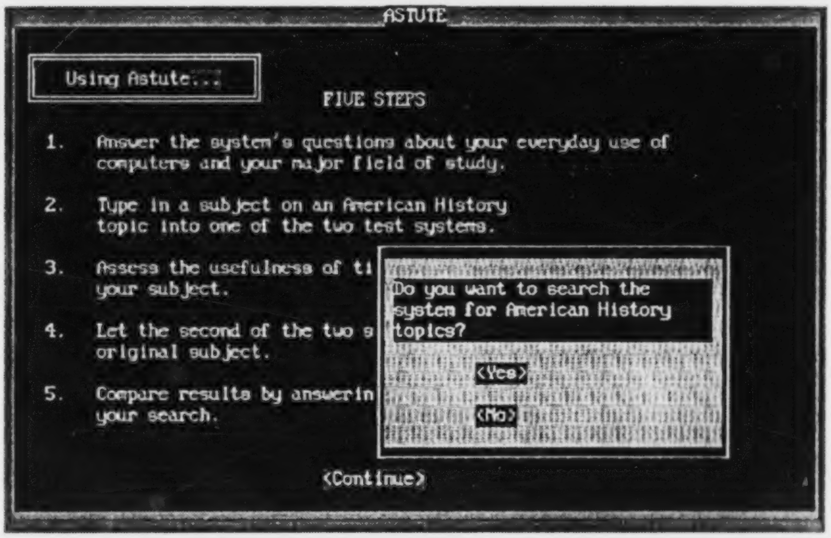
FIVE STEPS

1. Answer the system's questions about your everyday use of computers and your major field of study.
2. Type in a subject on an American History topic into one of the two test systems.
3. Assess the usefulness of titles the test system finds on your subject.
4. Let the second of the two systems try to find titles on your original subject.
5. Compare results by answering the system's questions about your search.

Figure 6.5. Fifth introductory screen

When users clicked on the <Continue> button in the fifth introductory screen, ASTUTE responded with a pop-up window asking users whether they wanted to search the system for engineering or computer science topics (at UM-Dearborn) or American history topics (at Earlham). The answer to this question gave users who did not want to search ASTUTE a graceful way of exiting the system. That is, the system took users who clicked on the <No> button back to the introductory screen. Users could also terminate their use of the system at any time by just walking away. The system responded to users who clicked on the <Yes> button with the first of three pre-search questions. Figure 6.6 shows the pop-up screen that is the system's response to UM-Dearborn users who clicked on the <Continue> button in the fifth introductory screen.

BEST COPY AVAILABLE



ASTUTE

Using Astute...

FIVE STEPS

1. Answer the system's questions about your everyday use of computers and your major field of study.
2. Type in a subject on an American History topic into one of the two test systems.
3. Assess the usefulness of ti
   your subject.
4. Let the second of the two s
   original subject.
5. Compare results by answerin
   your search.

Do you want to search the system for American History topics?

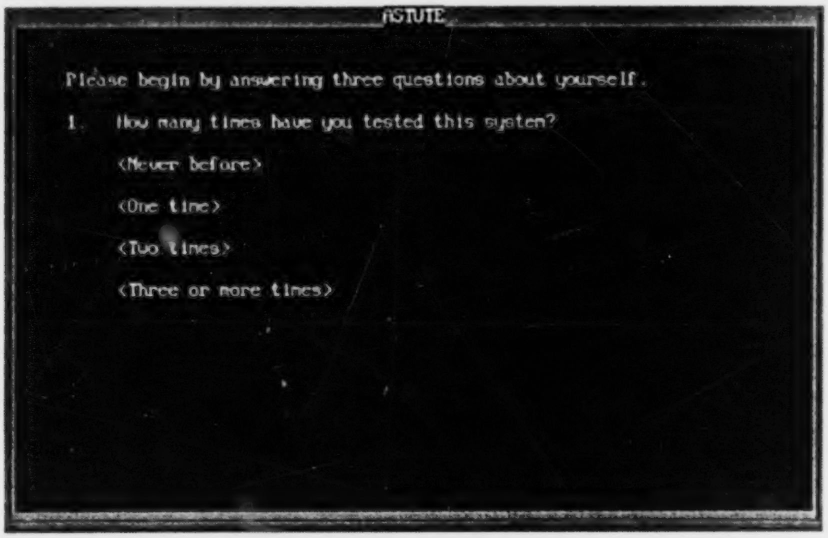
Figure 6.6. Pop-up window with question on test participation

6.4.2 Administering the Pre-search Questionnaire

When users clicked on the <Yes> button in response to the pop-up window's question on searching for topics in a particular subject, ASTUTE began recording all subsequent user actions and system responses to a transaction log.

The pre-search questionnaire contained three questions. Figures 6.7–6.9 show these three questions. Users could use the mouse to highlight and click on a response category or they could use the arrow keys on the keyboard to highlight the desired response category and press on the <Enter> key to select it. The first pre-search question asked users how many times they had used ASTUTE (figure 6.7).

BEST COPY AVAILABLE



Please begin by answering three questions about yourself.

1. How many times have you tested this system?

<Never before>

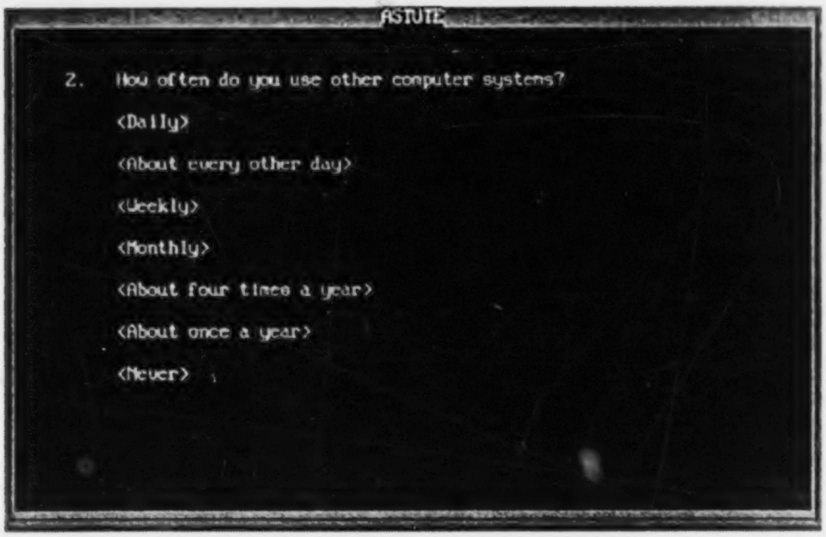
<One time>

<Two times>

<Three or more times>

Figure 6.7. First pre-search question on previous ASTUTE use

The second pre-search question asked users how often they used computer systems besides ASTUTE (figure 6.8).



2. How often do you use other computer systems?

<Daily>

<About every other day>

<Weekly>

<Monthly>

<About four times a year>

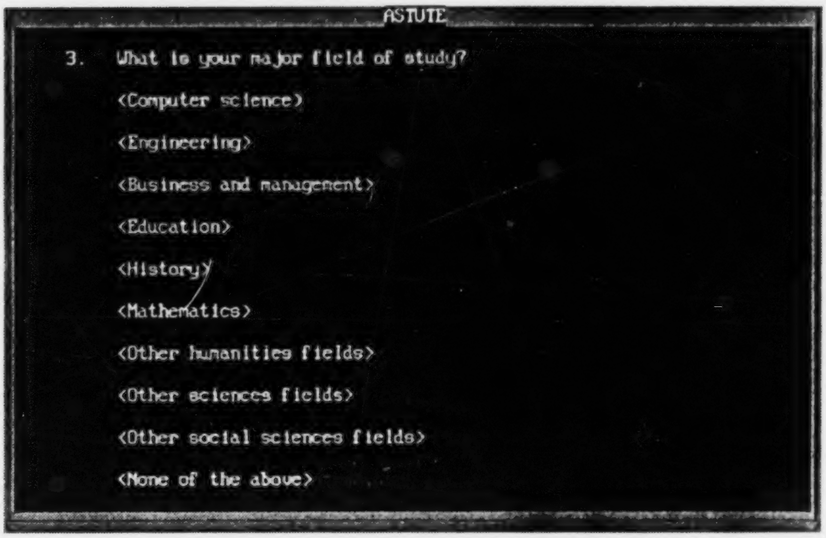
<About once a year>

<Never>

Figure 6.8. Second pre-search question on using other computer systems

The third pre-search question asked users to identify their major field of study (figure 6.9).

BEST COPY AVAILABLE



3. What is your major field of study?

<Computer science>

<Engineering>

<Business and management>

<Education>

<History>

<Mathematics>

<Other humanities fields>

<Other sciences fields>

<Other social sciences fields>

<None of the above>

Figure 6.9. Third pre-search question on major field of study

When ASTUTE began writing user actions and system responses to a transaction log, it activated a timeout function to determine whether users had terminated their search by walking away from the system. If ASTUTE recorded no user activity for four minutes, the system displayed a pop-up window asking the user to click on a <Continue> button to continue searching. If ASTUTE recorded no user activity within thirty seconds, the system closed the transaction log for the particular search administration, prepared to open the log for the next search administration, and displayed the first introductory screen (figure 6.1).

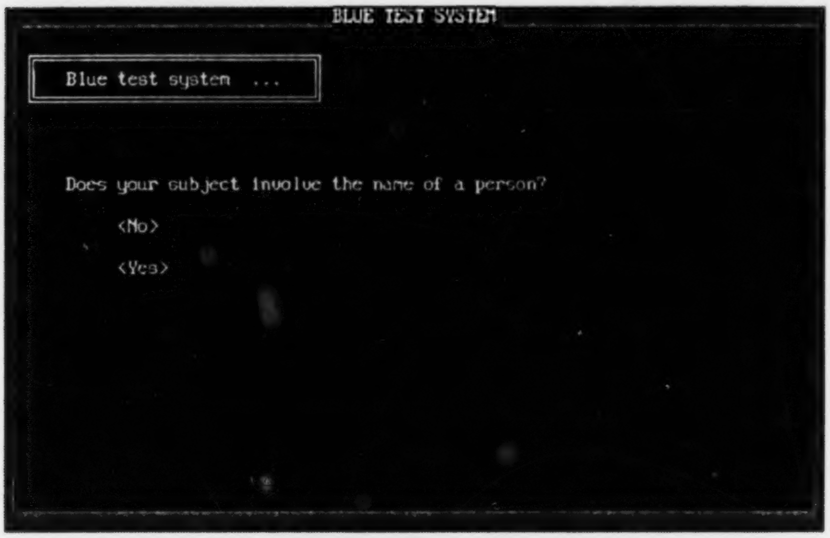
6.4.3 Searching the Experimental Online Catalog

The Comparison Search Experiment was designed to be as realistic as possible in that library patrons were searching the experimental online catalog with their own search topics and assessing the usefulness of retrieved items. Furthermore, transaction logging and online administration of pre- and post-search questionnaires made the data collection process as unobtrusive as possible.

When users clicked on <Yes> to the question in the pop-up window on the fifth introductory screen (figure 6.6), ASTUTE began logging user activity and system responses. It also assigned each participant to one of two odd/even experimental conditions. Odd-numbered participants searched the Blue Test System first, and even-numbered participants searched the Pinstripe Test System first.

BEST COPY AVAILABLE

Following the third pre-search question, the experimental online catalog asked users whether their query involved the name of a person. The Blue System used answers to this question to select between search trees for subjects generally and search trees for personal names as subjects. Figure 6.10 shows this question on personal names in subject queries.



```
BLUE TEST SYSTEM

Blue test system ...

Does your subject involve the name of a person?

     <No>

     <Yes>
```

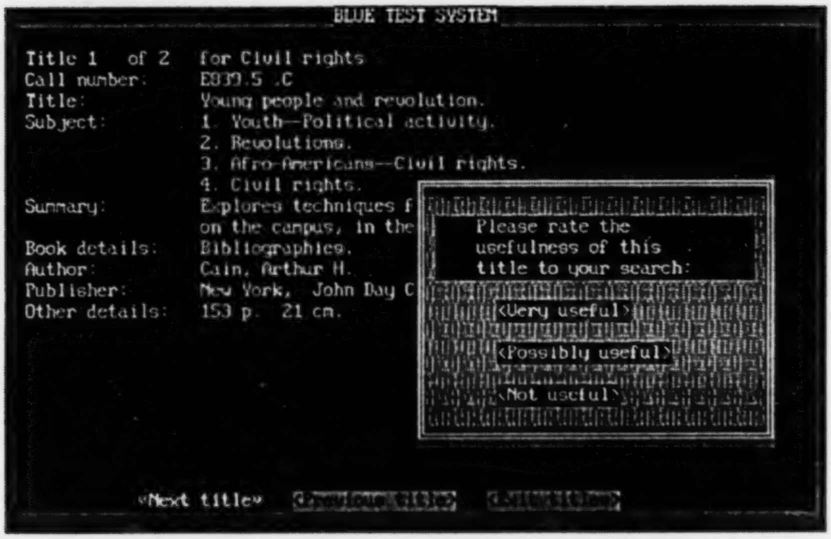
Figure 6.10. Question on personal names in subject queries

The Blue or Pinstripe System then asked users to enter their subject queries. (Chapter 5 details searching in the Blue and Pinstripe Systems.) As long as participants did not make a move to start a new search or enter a new query, they could search the Blue or Pinstripe System for as long as they wanted. When they made such a move, the Blue System (for odd-numbered participants) switched and automatically conducted a search for the original user-entered query in the Pinstripe System; the Pinstripe System (for even-numbered participants) switched and automatically conducted a search for the original user-entered query in the Blue System. As long as participants did not make a move to start a new search or enter a new query, they could search the second system for as long as they wanted. When they made such a move, the system initiated the post-search questionnaire. Both Blue and Pinstripe Systems recorded all user activity and system responses to transaction logs.

When users displayed bibliographic records, both Blue and Pinstripe Systems asked users to assess their usefulness. In figure 6.11, the Blue System displays a bibliographic record retrieved in a search for "civil rights" bearing a pop-up window that gives users three categories for rating usefulness: (1) useful, (2) possibly useful, and (3) not useful.

BEST COPY AVAILABLE

The user could click on one of these three categories using the mouse or use a combination of arrow keys and <Enter> key to select a category. Both Blue and Pinstripe Test Systems recorded user assessments of retrieved records to transaction logs.



```
Title 1   of 2    for Civil rights
Call number:      E839.5 .C
Title:            Young people and revolution.
Subject:          1. Youth--Political activity.
                  2. Revolutions.
                  3. Afro-Americans--Civil rights.
                  4. Civil rights.
Summary:          Explores techniques f
                  on the campus, in the
Book details:     Bibliographies.
Author:           Cain, Arthur H.
Publisher:        New York,  John Day C
Other details:    153 p.  21 cm.

Please rate the
usefulness of this
title to your search:
```

Figure 6.11. Relevance assessment categories

6.4.4 Administering the Post-search Questionnaire

The post-search questionnaire contained eleven questions. Appendix C lists these questions. Questions 4–6 asked users to compare and assess the performance of the two systems in view of the useful titles they retrieved. Questions 7–12 compared subject searching capabilities and ease of use of the two systems and library computer catalogs generally. Questions 13–14 were general questions about the subjects users search in online catalogs. Users could use the mouse to highlight and click on a response category or they could use arrow keys on the keyboard to highlight desired response categories and press on the <Enter> key to select them. To answer question 12, users typed in numbers from 1 to 5; if they entered any character except a number from 1 to 5, the system displayed an error message and allowed users to re-enter a number. All questions were closed-ended except for question 14. If users responded "very interested" to a question about their interest in seeing ASTUTE's capabilities extended to topics besides the ones in ASTUTE, the system asked users to type in the subjects that interested them. ASTUTE logged user responses to post-search questions to a transaction log.

BEST COPY AVAILABLE

6.5 Comparison Search Experiment (Library Staff)

6.5.1 Background

In preparation for submitting the proposal to the Department of Education, the project director approached UM-Dearborn and Earlham College library staff about participating in the project. When the project director described search trees to UM-Dearborn staff, they bristled at the idea of an online catalog selecting a subject searching approach on its own in response to their subject queries. They suggested that the project director extend the Comparison Search Experiment to library reference and cataloging staff to obtain their reactions to the experimental system. Thus, the project director added a Search Experiment to the evaluation section of the Department of Education proposal that included library staff.

6.5.2 Recruiting Library Staff

The project director had informed UM-Dearborn and Earlham College library staff of their participation in the Comparison Search Experiment in written correspondence, proposal drafts, and face-to-face discussions before submitting the proposal to the Department of Education. In fact, the suggestion to include library staff in the search experiment came from UM-Dearborn library staff.

While the experimental online catalog was available for patron use at UM-Dearborn and Earlham, the project director sent staff letters reminding them of and describing the search experiment and asking for their participation. Library staff liaisons suggested the names of prospective participants, scheduled one-hour interviews with those who volunteered their participation, or the project director called them directly and scheduled them herself. The project director sent staff participants the cover letter and consent form given in Appendix D. The cover letter asks library staff to bring three to five subject queries in the area of computer science or engineering (at UM-Dearborn) or American history (at Earlham) to search. The queries could come from library patrons in the course of their reference work or could be subject queries in which they had a professional or personal interest. The cover letter also assured volunteers of confidentiality.

Interviews were scheduled on the day(s) immediately preceding and on the day that the ASTUTE project team disassembled the microcomputer equipment and transported the system back to Ann Arbor. This gave library staff several weeks to search ASTUTE on their own and saved the project team an extra trip to participating libraries.

6.5.3 Comparison Search Experiment Procedures

At the scheduled time, the project director and library staff member met where ASTUTE was installed and began the Comparison Search Experiment. Assuming the role of interviewer, the project director introduced herself to the staff member, gave a brief overview of the project, and described how the search experiment would proceed.

Library staff participants searched ASTUTE for one of the topics they brought. They responded to the three questions in the pre-search questionnaire, conducted their searches on the Blue and Pinstripe Systems, and responded to the eleven questions in the post-search questionnaire. The interviewer then asked participants nine open-ended questions about the impetus for their queries, difficulties searching the test systems, suggestions for system improvements, and system preference. She wrote their answers in spaces provided on the Post-search Interview Form with Library Staff (appendix E).

Participants then searched ASTUTE for a second time and responded to ASTUTE's pre- and post-search questionnaires. The interviewer asked participants the same nine open-ended questions about their searching experiences. Participants then searched ASTUTE for third and last time, responded to ASTUTE's pre- and post-search questionnaires, and the interviewer asked participants the same nine open-ended questions about their searching experiences. Following these nine questions, the interviewer asked four summary questions to elicit library staff comments on what they liked most and least about the experimental online catalog, their overall system preference, and any additional comments staff wanted to make.

Each library staff participant was randomly assigned to one of two odd/even experimental conditions. Odd-numbered participants first searched the Blue Test System, and even-numbered participants first searched the Pinstripe Test System. Occasionally, library staff entered queries on computer science/engineering topics (at UM-Dearborn) or American history topics (at Earlham) which were classified in different areas from the Library of Congress Classification areas selected for inclusion in ASTUTE databases. In such instances, the interviewer disregarded the query and allowed participants to conduct a search for another query.

Appendix E gives the questionnaire used by the interviewer in the Comparison Search Experiment with library staff. The questionnaire contains the interviewer's introductory statement, open-ended questions following each search in the Blue and Pinstripe Systems, and summary questions following searches.

6.6 Recording Search Experiments to a Transaction Log

ASTUTE recorded individual administrations of the Comparison Search Experiment to a transaction log. The log consisted of up to four record levels for each patron or staff participant in the Comparison Search Experiment. Records at all four levels contained a unique identification number to enable the researchers to link the various records to one another. Level 1 records contained search experiment beginning and ending time stamps and answers to pre- and post-search questionnaires. Level 2 records documented user actions and system responses. Level 3 records contained relevance assessments for displayed bibliographic records. Level 4 records contained user responses to the only open-ended question in the questionnaire on user interest in seeing ASTUTE's capabilities extended to topics besides the ones in ASTUTE. Appendix F details the content and format of the four levels of transaction log records in the experimental online catalog.

6.7 Timeout Function

The ASTUTE project team did not monitor ASTUTE use at the two participating libraries. Thus, users could come and go and use the system for as short or as long as they wanted. We added a timeout function to ASTUTE because there was nothing to stop users from walking away from the system at any time. FoxPro had a timeout function and we used this function to reset the program back to the initial screen after a period of time had lapsed without user activity. Originally, we set the timeout function to three minutes. During the pretest, the timeout function terminated several users' searches prematurely. We tried longer timeout periods and finally decided on a four-minute timeout period. If ASTUTE did not detect user activity for four minutes, it asked users the following question, "If you are still participating in the test, please press the <Enter> key or press the left mouse button to <Continue>." ASTUTE gave users thirty seconds to respond to this question. If users did not respond, the system recorded the time in level 1 transaction log records, terminated the search, and displayed the first introductory screen (see figure 6.1).

6.8 Design of the Comparison Search Experiment

6.8.1 Test Administration

The project director chose the Comparison Search Experiment because it featured a partially controlled but authentic experiment. Respondents at the two participating libraries conducted online searches on a topic of their own choosing — sequentially —

on Blue and Pinstripe Test Systems, judged the relevance of retrieved records, and answered questions about their searching experiences in the two systems.

The Comparison Search Experiment that enlisted library patrons used the same design as Comparison Search Experiments conducted by library patrons in previous studies (Siegel et al. 1983; Markey and Demeyer 1986. 109 ff.). The Comparison Search Experiment involving library staff deviated from the experiment with library patrons as follows: (1) staff searched the Blue and Pinstripe Systems for three different topics, (2) staff took part in an extended post-search interview that featured open-ended answers to questions on their system use and experiences, and (3) staff conducted subject searches for topics that might have originated from sources other than their own personal experiences and interests.

Table 6.1 lists dependent and independent variables measured in the Comparison Search Experiments with library patrons and staff.

Table 6.1. Variables Measured in Comparison Search Experiment

Dependent Variables:
Number of retrieved bibliographic records
Number of displayed bibliographic records
Number of useful records displayed
Time spent searching Blue and Pinstripe Test Systems
System capabilities used
Ease of system and system capability use
Satisfaction with subject searches
System preference
Independent Variables:
Type of user (patron or staff)
System used (Blue or Pinstripe)
First system searched
Subject area (per library)
Subject searching approaches to be executed

Table 6.2 summarizes the procedures used in Comparison Search Experiments with library patrons and staff.

Table 6.2. Comparison Search Experiment Procedures Summary

Procedures	Patron experiment	Staff experiment
Source of search topic	Patron	Staff member
System searcher	Patron	Staff member
Source of relevance assessments	Patron	Staff member
Interviewer	Experimental system	Project director and experimental system
Estimated completion time	15–20 minutes	60 to 75 minutes

In the Comparison Search Experiment with patrons, two systems (Blue and Pinstripe) were studied at the two participating libraries. The system that was first searched by the patron was varied so that approximately half of the participants searched the Blue System first and the other half first searched the Pinstripe System. Since monitors were not present to encourage patrons to complete the entire Comparison Search Experiment, patrons could leave the experiment at any time. The researchers expected full and partial administrations of the Comparison Search Experiment. Table 6.3 summarizes the events completed in full and partial administrations of this experiment. Codes given in the bottom row of this table designate whether the collected data were considered full or partial usable administrations administrations.

Table 6.3. Full and Partial Administrations of the Comparison Search Experiment

Events	Completed events				
Pre-search questionnaire responses	x	x	x	x	x
Pinstripe System search	x		x	x	x
Blue System search		x	x	x	x
Post-search questionnaire responses					x
Inconsistent post-search questionnaire responses				x	
Disposition of test administration	u, p	u, p	u, p	u, p	u, f

Key to disposition:
u = usable test
p = partial test
f = full test

There were four conditions under which usable, partial test administrations were possible. Full test administrations involved completion of pre- and post-search questionnaires, and searches in Blue and Pinstripe Systems. Post-search questionnaire responses in full test administrations had to be consistent with other data such as test system searches and relevance assessments.

Table 6.3 does not cite the characteristics of unusable test administrations. The project team expected several conditions under which unusable test administrations would occur, e.g., entry of queries on subjects other than computer science or engineering (at UM-Dearborn) or American history (at Earlham), occurrence of system errors. Several sections of chapter 7 detail characteristics of unusable test administrations.

Monitors were present to encourage library staff to complete the entire Comparison Search Experiment. Administrations of the Comparison Search Experiment included responses to pre- and post-search questionnaires, searches in Blue and Pinstripe Test Systems, and the staff member's answers to open-ended questions posed by the interviewer in the extended post-search interview. Except for the unfortunate occurrence of system errors that might have disrupted administrations of the Comparison Search Experiment with library staff, the project team expected to collect only full administrations of this experiment.

6.8.2 Quantitative Data Analysis

The experimental online catalog's transaction logging capability recorded subject searches, relevance assessments, and questionnaire responses. Subject searches provided quantitative data on the following:

- Length of user queries.
- Time spent searching.
- Frequency of system capability use.
- Number of titles retrieved.
- Number of titles displayed.

The combination of these data with user relevance assessments enabled the researchers to compare the performance of the Blue and Pinstripe Systems in terms of:

- Precision (defined as the proportion of retrieved documents that users rate "very useful").
- Estimated recall (defined as the fraction of titles users rate "very useful" in relation to all titles users rate "very useful" in the course of their searches for their topics of interest in both Blue and Pinstripe Systems). This

definition of estimated recall was adapted from a landmark study of information seeking and retrieving (Saracevic and Kantor 1988).

The researchers used t-test procedures to compare mean scores of precision and estimated recall in the Blue and Pinstripe Systems and to ascertain whether differences in mean scores were statistically significant (see chapter 9). We supplemented the results of statistical procedures with user responses to questions in the post-search questionnaire on ease of system and system capability use, user satisfaction with subject searches, and user system preference.

6.8.3 Qualitative Data Analysis

The failure analysis of Blue- and Pinstripe-system searches used a simple measure to characterize search success and failure. A successful search was one in which users retrieved at least one title that they rated "very useful." A failed search was one in which users failed to retrieve at least one title that they rated "very useful." The project director, an expert searcher of information retrieval systems, repeated successful and failed searches to determine the reasons why searches succeeded or failed.

The failure analysis featured an investigation of failed searches (see chapter 10). It also focused on failed and successful searches per subject searching approach type (see chapter 11). The latter identified problems with the implementation of a particular subject searching approach and needed improvements. It also demonstrated successful implementations of subject searching approaches that required little or no fine tuning.

Open-ended questions were limited to the extended post-search questionnaire administered by an interviewer to library staff. The interviewer asked staff nine questions about their subject searching experiences in the Blue and Pinstripe Test Systems after they had performed subject searches in both systems on a topic of their own choosing.

1. What was the impetus for the subject you searched?
2. What difficulties did you have searching for your subject in the Blue and Pinstripe Systems?
3. What improvements could be made to the Blue and/or Pinstripe Systems?
4. What differences did you notice between the ways in which the two systems responded to the subjects you typed in?
5. In what ways did the Blue and/or Pinstripe Systems give you new ideas for searching for subjects?

6. What differences did you notice between the titles the Blue and Pinstripe Systems found for your subject?
7. What did you like most and least about Blue and/or Pinstripe Test Systems?
8. What is your overall system preference and why?

Answers to these questions were not always on the tip of the staff member's tongue and had to be drawn out through the interviewer's probing. This probing helped to focus the staff member's attention on a specific Blue or Pinstripe System capability and to elicit their kudos about the system(s) as a whole or specific system capabilities and their suggestions for improving the Blue and Pinstripe Systems. (Probes are listed in open-ended questions in the extended post-search questionnaire in appendix E.) In two separate questions, staff were asked what they liked most and least about the Blue and Pinstripe Test Systems. They were also asked which system they preferred overall and why.

6.9 Chapter Summary

This chapter featured a discussion of data collection procedures and instruments used in the evaluation of the performance of the experimental online catalog (see sections 6.4–6.5).

The researchers transported and assembled a Gateway microcomputer bearing ASTUTE to the two data collection sites — Mardigian Library at the University of Michigan-Dearborn and Lilly Library at Earlham College. The microcomputer was dedicated to use of the ASTUTE experimental online catalog. The ASTUTE experimental online catalog performed recruiting functions on its own. The presence of the microcomputer equipment in a public place in the library, posted written signs, and two alternating introductory screen savers attracted potential respondents.

This study enlisted the Comparison Search Experiment because it featured a partially controlled but authentic experiment (Siegel et al. 1983). Respondents at the two participating libraries conducted online searches on a topic of their own choosing — sequentially — on Blue and Pinstripe Systems, judged the relevance of retrieved records, and answered questions about their searching experiences in the two systems. When users displayed bibliographic records, both Blue and Pinstripe Systems asked users to assess their usefulness. Transaction logging and online administration of pre- and post-search questionnaires made the data collection process as unobtrusive as possible (see section 6.6 and Appendix F). Qualitative and quantitative analyses of collected data were described (see section 6.7).

References

Belkin, Nicholas J. et al. 1990. "Taking account of users tasks, goals and behavior for the design of online public access catalogs." In *ASIS '90; proceedings of the 53rd ASIS annual meeting; 4–8 November, 1990, Toronto*, edited by Diane Henderson, 69–79. Medford, N.J.: Learned Information.

Hancock-Beaulieu, Micheline, Lorna McKenzie, and Avril Irving. 1991. *Evaluative protocols for searching behaviour in online library catalogues*. London: British Library. British Library R&D Report 6031.

Hildreth, Charles R. 1993. *An evaluation of structured navigation for subject searching in online catalogues*. Ph.D. dissertation. Department of Information Science, The City University, London.

Jones, Richard M. 1988. *A comparative evaluation of two online public access catalogues*. London: British Library. British Library Research Paper 39.

Markey, Karen, and Anh N. Demeyer. 1986. *Dewey Decimal Classification online project: Evaluation of a library schedule and index integrated into the subject searching capabilities of an online catalog*. Dublin, Ohio: OCLC. OCLC Research Report OCLC/OPR/RR–86/1.

Saracevic, Tefko, and Paul Kantor. 1988. "A study of information seeking and retrieving." *Journal of the American Society for Information Science* 39, 3: 197–216.

Siegel, Elliott R. et al. 1983. "Research strategy and methods used to conduct a comparative evaluation of two prototype online catalog systems." In *National Online Meeting proceedings —1983*, 1983 April 12–14, New York, compiled by Martha E. Williams and Thomas H. Hogan, 503–11. Medford, N.J.: Learned Information.

University of California, Division of Library Automation and Library Research and Analysis Group. 1982. *Users look at online catalogs: results of a national survey of users and non-users of online public access catalogs: final report to the Council on Library Resources*. Berkeley, Calif.: University of California, November 16.

Walker, Stephen, and Richard M. Jones. 1987. *Improving subject retrieval in online catalogues; 1. Stemming, automatic spelling correction and cross-reference tables*. London: British Library. British Library Research Paper 24.

Walker, Stephen, and Rachel De Vere. 1990. *Improving subject retrieval in online catalogues; 2. Relevance feedback and query expansion*. London: British Library. British Library Research Paper 72.

7 Comparison Search Experiment Participants

7.1 Introduction

Although previous researchers had enlisted the Comparison Search Experiment to compare and evaluate computerized library systems, the experiment conducted in this study had three unique aspects: (1) the system recruited library patrons, (2) interviewed them, and (3) recorded their searches and interview responses. Since monitors were not present to encourage patrons to complete the entire Comparison Search Experiment, the researchers expected patrons to leave before they finished the full administration of the Comparison Search Experiment (see Table 6.3).

Before the researchers could submit Comparison Search administrations to the data analysis, they had to manually review individual administrations to determine whether patron search topics were in scope and whether patron responses to pre- and post-search questionnaires were valid. This chapter describes the results of this manual review process.

7.2 Michigan Pretest

Prior to the installation of the experimental online catalog in the two participating libraries, the project director conducted a pretest of the Comparison Search Experiment. She recruited 13 students enrolled in the School of Information and Library Studies at the University of Michigan to take part in the pretest. In interviews lasting one hour, the project director observed students and encouraged them to comment on all aspects of the experiment, e.g., user interface, wording of pre- and post-search questions, system functionality. On the basis of the pretest at the University of Michigan, the project team made several changes to the experimental system and to Comparison Search Experiment procedures. Examples of changes are:

- Changed wording of system options, instructions, explanatory messages, and questions and responses on pre- and post-search questionnaires.
- Added response categories to pre- and post-search questionnaires.
- Changed colors of highlighted options.
- Increased timeouts.
- Disabled certain keys on the keyboard.
- Identified and fixed bugs.

7.3 Installation Period Length

Following the Michigan pretest, the researchers transported and assembled the Gateway microcomputer bearing ASTUTE to the first of two data collection sites — Mardigian Library at the University of Michigan-Dearborn. The system was available to library patrons for a five-week period beginning on November 11, 1993. The Mardigian Library staff liaison instructed evening and weekend staff in system start up, shut down, and backup procedures. The liaison and other library staff took the responsibility of daily system maintenance, i.e., turning the system on and off, making periodic backups of transaction log files. Periodically, the ASTUTE project team would visit Mardigian Library to make sure that the system was functioning properly and make backups of transaction log files.

The project director conducted the Comparison Search Experiment with Mardigian Library staff in the fifth week of system installation. When the project director conducted the first of four interviews with staff, the system unexpectedly crashed. The systems programmer studied the problem and found system and transaction log files had been corrupted when users unexpectedly turned system equipment on and off.

ASTUTE project staff studied logged searches to determine the extent of the damage. They decided that the number of corrupted searches was so large that they doubted the reliability of collected data. They informed the Mardigian Library staff liaison of the problem who invited the project team back to the library for a second installation and data collection period.

Investigating the problem of corrupted system and log files, the ASTUTE project team introduced the following changes to ASTUTE to ensure that files would not be corrupted in future system installations at participating libraries:

- The main problem was that the <Reset> button on the front of the computer enticed users to push it. This sometimes corrupted the database tables. To solve the problem, we disconnected the wire that connected the <Reset> button to the computer.
- We programmed the autoexec.bat file to automatically restart the ASTUTE experimental online catalog in the event that the machine was restarted or rebooted. This discouraged students from using the machine for other purposes.
- ASTUTE project team members personally monitored the log tables frequently to be sure the machine was in use.
- Checks were also made to ensure that the computer equipment and mouse were functional.
- When several track balls disappeared from the mice at participating libraries, we replaced non-functioning mice with inexpensive mice.
- We added a password protection feature to exit the ASTUTE experimental online catalog. This discouraged students from using the machine for other purposes.

The researchers transported and assembled the Gateway microcomputer bearing the updated ASTUTE to the second of two data collection sites — Lilly Library at Earlham College. The system was available to library patrons for a fifteen-week period beginning on February 23, 1993. The Lilly Library staff liaison instructed evening and weekend staff in system start up, shut down, and backup procedures. The liaison and other library staff assumed the responsibility for daily system maintenance, i.e., turning the system on and off, making periodic backups of transaction log files. In lieu of periodic visits to this distant site to make sure that the system was functioning properly and make backups of transaction log record files, the Lilly Library liaison sent transaction log records to the ASTUTE project team via file transfer protocol (ftp) or ground mail. The ASTUTE project team reviewed the logs as soon as possible to make sure everything was functioning properly.

ASTUTE was installed in Lilly Library for a much longer time period than was given in the original schedule in the project proposal. Lilly Library staff advised the project director to extend the system installation period for several reasons: (1) Earlham College had far fewer students than UM-Dearborn (1,100 vs. 7,400 students), (2) the ASTUTE project team had installed ASTUTE in Lilly Library during the second half of the College's winter quarter and library staff had not had the opportunity to inform students of the system's availability, and (3) keeping ASTUTE installed in Lilly Library through the College's spring semester would give library staff the opportunity to tell students about the system in the many classes staff teach and in their daily interactions with students.

Having installed ASTUTE in Lilly Library, the ASTUTE project team turned their attention to making changes to the ASTUTE experimental online catalog to ensure a successful second administration at Mardigian Library. The team installed ASTUTE in Mardigian Library on March 12, 1993 and left it there for five weeks. This time period coincided with the second half of the university's thirteen-week winter semester. The ASTUTE project team regularly visited Mardigian Library to make sure that the system was functioning properly and to make backups of transaction log record files. The Mardigian Library staff liaison sent transaction log records to the ASTUTE project team via two-day campus mail. The ASTUTE project team reviewed the logs as soon as possible to make sure everything was functioning properly.

Table 7.1 gives the dates of system installation in the two libraries and number of administrations of the Comparison Search Experiment with library patrons and staff.

Table 7.1. System Installation Information

Location	Date	Search Administrations*
Mardigian Library	Nov. 11–Dec. 14, 1992	556
Mardigian Library	Mar. 12–Apr. 19, 1993	826
Lilly Library	Feb. 23–May 28, 1993	238

*Includes library patron and staff searches.

Sections 7.4–7.6 describe the characteristics of usable and unusable searches from the three data collection periods.

7.4 First Comparison Search Experiment Administration at UM-Dearborn

The first data collection period at UM-Dearborn lasted five weeks from November 11, 1992 to December 14, 1992. The experimental online catalog administered a total of 556 search experiments. The log includes twelve administrations of the Comparison Search Experiment with library staff. The researchers manually reviewed the queries entered in these experiments to determine usable administrations for submission to data analyses. Figure 7.1 divides search administrations into five components for "usable" and "unusable" search administrations.

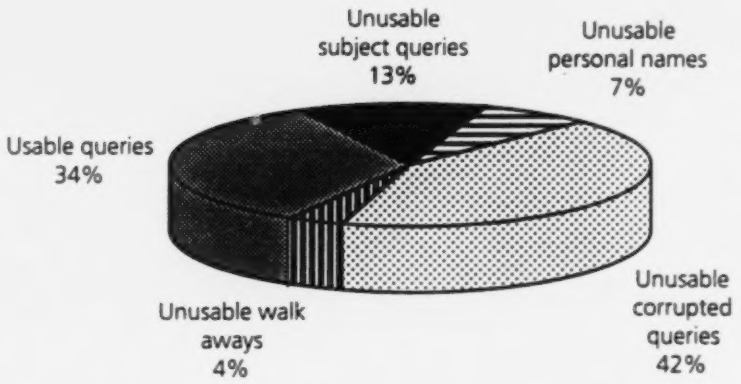


Figure 7.1. Usable and unusable administrations of the first Comparison Search Experiment at UM-Dearborn

Reviewing logged data, the ASTUTE project team determined that experimental data logged during the one week preceding and one week following Thanksgiving had been corrupted by the user action of repeatedly turning the microcomputer on and off. The corruption affected 234 searches. Thus, a large number of searches in the middle of the log were unusable. Since the team was concerned about the overall reliability of the data, they disregarded data from the first collection period and planned a second data collection period at UM-Dearborn. The only data they kept were the twelve administrations of the Comparison Search Experiment with library staff.

The ASTUTE project team treated data from this first data collection period as a pretest. The team introduced changes to the experimental online catalog to ensure that data files would not become corrupted in future installations of the catalog in participating libraries and disconnected certain keys on the keyboard.

Search experiments in the "walk-aways" category in figure 7.1 accounted for 4% (22 of 556) of UM-D search administrations. These were search administrations in which users completed one or more pre-search questions but they did not continue with their searches. These users probably walked away from the system and it eventually reset itself to the introductory screen savers.

Unusable administrations accounted for 20% of first-comparison administrations. Unusable administrations of the Comparison Search Experiment were connected with the content of user queries. Such queries probably would have failed to produce useful retrievals in the experimental online catalog because they were out of scope, that is, they did not match the subject coverage of the catalog's database at UM-Dearborn (i.e., computer science and technology).

"Usable" administrations of the Comparison Search Experiment accounted for about one-third of the queries logged in the first data collection period at UM-Dearborn. Over half (58%) were full administrations of the search experiment. A large percentage (26%) of "usable" search administrations contained all events except a completed post-search questionnaire.

7.5 Second Comparison Search Experiment Administration at UM-Dearborn

7.5.1 Introduction

The second data collection period at UM-Dearborn lasted five weeks from March 12, 1993 to April 19, 1993. The experimental online catalog administered a total of 826 search experiments. The researchers manually reviewed the queries entered in these experiments to determine usable administrations for submission to data analyses. Subsections below describe the characteristics of usable and unusable search experiments. Figure 7.2 divides search administrations into four components. The largest component was for "usable" search administrations for subject queries.

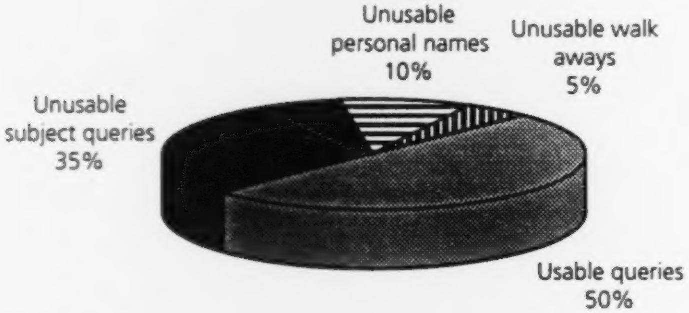


Figure 7.2. Usable and unusable administrations of the second Comparison Search Experiment at UM-Dearborn

Of the 826 search administrations, about half (412 of 826 administrations) were usable. Over a third (35%) were unusable queries that users entered using the experimental online catalog's subject searching capabilities for subjects generally. About 10% were unusable queries entered using searching capabilities for personal names. The remainder (5%) were search administrations in which users completed one or more pre-search questions but they did not continue with their searches. These users probably walked away from the system and it eventually reset itself to the introductory screen savers.

Figure 7.3 depicts daily activity in terms of the total number of search administrations and usable search administrations.

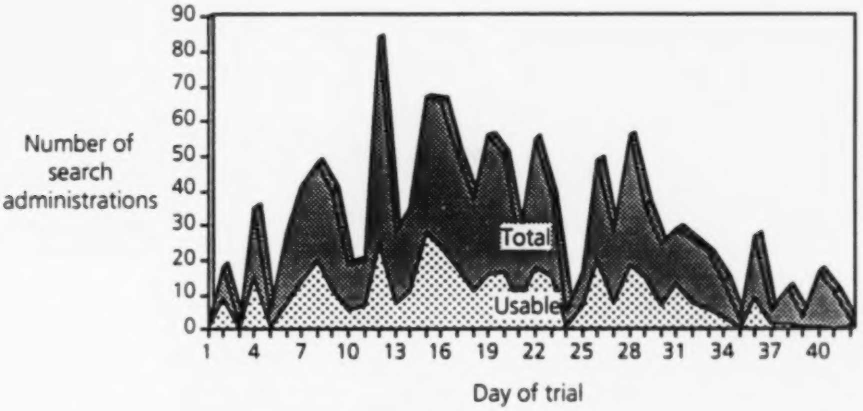


Figure 7.3. Search administrations at UM-D on a daily basis

An increase or decrease in the number of usable administrations was accompanied by a corresponding increase or decrease in the total number of unusable administrations. Toward the end of the data collection period, the total number of unusable administrations increased without a concomitant increase in usable administrations. Reviewing the logs, the researchers discovered a key-remapping error that had been introduced into the system on the thirty-fourth day of data collection. When users typed the letter "r," the system understood it as the letter "i." Thus, the researchers placed several administrations with queries such as "etheinet" and "image piocessing" in the unusable category called "system error." These administrations would have been usable had the keys not been remapped. This might help explain why there were many unusable administrations toward the end of the data collection period.

Figure 7.4 shows daily activity by day of the week. Searching was heaviest in the early part of the week (Tuesdays, Wednesdays, Thursdays). It decreased toward the early part of the weekend and was heavy again on Sundays. In a related study, Neal Kaske

(1988, 366) found subject searching especially heavy at the Engineering branch library at the University of Alabama on Wednesdays and weekends.

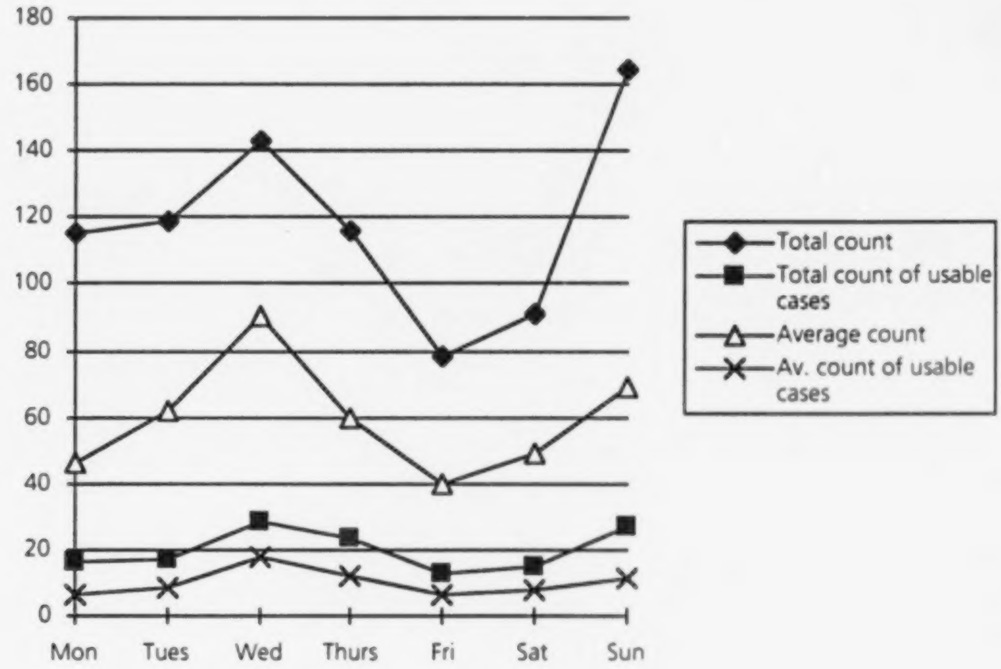


Figure 7.4. Usable search administrations at UM-D by day of the week

7.5.2 Usable Administrations

Table 7.2 summarizes the characteristics of the 412 usable administrations in the second Comparison Search Experiment at UM-Dearborn.

Table 7.2. Usable Administrations in the Second Comparison Search Experiment at UM-D

Events	**Completed events**				
Pre-search questionnaire responses	x	x	x	x	x
Pinstripe System search	x		x	x	x
Blue System search		x	x	x	x
Post-search questionnaire responses					x
Suspect post-search questionnaire responses				x	
Number of usable search administrations (N=412)	59 (14%)	56 (14%)	55 (13%)	71 (17%)	171 (42%)

A large percentage (42%) of usable administrations of the Comparison Search Experiment were full administrations. Of the four partial administration categories, the largest percentage (17%) contained the four complete events; however, responses to the post-search questionnaire were suspect. That is, post-search questionnaire responses were default responses that users could have selected by repeatedly pressing the <Enter> key or by repeatedly entering the same number. In the data analysis, the researchers ignored suspect post-questionnaire responses and combined these administrations with search administrations in which pre-search questionnaires and searches in Blue and Pinstripe Systems had been completed. The latter administrations accounted for 13% of usable search administrations. When they were combined with the former, the two categories of usable search administrations amounted to almost a third (30%) of usable search administrations. Remaining usable administrations had searches in only one of the two test systems (Blue or Pinstripe) and each of these two categories amounted to 14% of usable administrations.

A total of four librarians took part in the Comparison Search Experiment at UM-D. These librarians contributed twelve full search administrations; however, searches for two full administrations were discarded because of a system error. This left ten full search administrations.

Chapters 8–12 feature the analysis of usable Comparison Search Experiment administrations.

7.5.3 Unusable Administrations

The second administration of the Comparison Search Experiment at UM-Dearborn yielded 414 unusable search administrations. Forty search administrations were unusable because users completed one or more pre-search questions but they did not enter a query into the catalog. These users probably walked away from the experimental online catalog. After four minutes had elapsed, system's timeout function reset itself to the initial introductory screen savers

Table 7.3 summarizes the characteristics of unusable administrations in which users entered queries into the experimental online catalog. The queries in these unusable administrations were entered through subject searching capabilities or through personal-name subject searching capabilities. The system's choice between the two types of capabilities was based on user responses to the question following the pre-search questionnaire that asked users whether their query involved the name of a person.

Table 7.3. Unusable Administrations of the Second Comparison Search Experiment at UM-D

	Capabilities for Subjects		**Capabilities for Personal Names**	
Query categories	No.	%	No.	%
Out of scope				
Subjects	128	44.1	1	1.2
Personal names	10	3.5	21	25.0
Names and subjects	0	0.0	9	10.7
Playing				
Subjects	17	5.9	0	0.0
Personal names	3	1.0	2	2.4
Names and subjects	2	0.7	10	11.9
Title	2	0.7	0	0.0
Meaningless input:				
Blank	69	23.8	1	1.2
Character	1	0.3	0	0.0
Gibberish	19	6.6	9	10.7
Letter	8	2.8	0	0.0
Letters	6	2.1	0	0.0
Known-item search				
Title search	3	1.0	0	0.0
Author-title	0	0.0	23	27.3
Author	1	0.3	2	2.4
Other unusable searches				
Sex term	10	3.4	3	3.6
Expletive	2	0.7	0	0.0
Command	1	0.3	0	0.0

Subject entered as a name	n/a	n/a	3	3.6
System error	8	2.8	0	0.0
Total	290	100.0	84	100.0

Most queries in unusable administrations were out of scope, that is, they did not match the subject coverage of the catalog's database at UM-Dearborn (i.e., computer science and engineering). Since the catalog was restricted to subject searches, queries requesting author and/or titles were also considered unusable.

7.5.3.1 Out-of-scope Subject Queries

Of the 290 queries entered through subject searching capabilities, the largest percentage (44.1%) were out of scope. Examples of these queries were "astrophysics," "croatia," "euthanasia," and "activity based costing."

Users averaged almost two access points per search. Two searchers entered as many as seven access points. One of the two searchers entered the query "interplationr." When the Pinstripe System informed the user that it was unable to find material and suggested that the user check the query's spelling, the user responded by clicking on the <Do search> six times before walking away from the experimental system. Thus, the system counted this access point seven times. The other user initially entered the query "saas." He responded to several Blue-system messages informing him that the system was unable to find material by entering new queries, e.g., "soft tissue" and "cancei" [sic].

Of the three subject searching capabilities available to the Pinstripe System for retrieving bibliographic records, the keyword-in-record search was most successful in making retrievals. The Blue System had seven subject searching capabilities available for retrieving bibliographic records. The alphabetical approach was the most successful. The exact approach was also responsible for retrieving bibliographic records. Examples of user queries which were successful in retrieving records through the exact approach were "physics," "differential equations," "trigonometry," and "foreign relations." Although users retrieved records for these and other queries, the researchers felt that the retrieved material did not focus entirely on their topics of interests and that users could retrieve a large quantity of material on such topics in catalog that was encyclopedic in coverage. Even though these queries retrieved bibliographic records, the researchers deemed them out of scope in a catalog limited to computer science and technology material.

Of out-of-scope subject searches, 42.2% of searches were conducted in one of the two experimental online catalogs. The average number of access points per search was comparable to out-of-scope subject searches conducted in both systems.

There was only one out-of-scope *subject query* entered through *personal-name subject searching capabilities*. The query was "medical statistics." The user searched only the Blue System which was unable to retrieve bibliographic records.

7.5.3.2 Out-of-scope Queries for Personal Names

Users entered out-of-scope personal name queries through both general subject searching and personal-name subject searching capabilities. Examples of out-of-scope queries for personal names that were entered using the system's subject searching capabilities were "assad," "bush, g," "plotkin," and "sinclair upton." Users entered about one query per search.

In the Blue System, search trees for personal-name searching invoked keyword approaches before responding to user queries with the alphabetical approach regardless of the extent to which user queries matched an indexed name. Users conducted only two searches for out-of-scope personal names in a single system — the Blue System. The two queries were "hausdorf" (last name) and "kumeso michael matabi" (last name, first name, topic). Both searches in this system gave users the opportunity to select personal-name subject headings through the alphabetical approach. The user in the former search selected the name "Land, Edwin Herbert," and displayed the one bibliographic record bearing this name. The user in the latter search selected the name "Ford, Henry," and displayed the eight bibliographic record bearing this name. Both users gave displayed records "very useful" ratings.

There were nineteen searches in which users searched both Blue and Pinstripe Systems. The Blue System gave users the opportunity to select personal-name subject headings through the alphabetical approach. Although listed names did not refer to the person in the user query, several users selected listed names, displayed bibliographic records, and gave displayed records "very useful" ratings. Examples of user queries and personal-name headings users selected were:

User query	Selected name heading
corniliustacitus	Salomon, Erich
pasteur, louis	Zeppelin, Ferdinand
tolstoy, leo	Adams, Ansel
seppalainen	Shute, Nevil
Zhao, X	Zeppelin, Ferdinand

Three queries processed by the Pinstripe System retrieved bibliographic records through the keyword-in-record approach for personal names. Surprised by these retrievals, ASTUTE project team members repeated these searches to determine why they produced retrievals. They analyzed the results and realized that they had erroneously included main- and added-entry fields for personal names (USMARC fields 100 and 700) when they had selected bibliographic record fields for the database supporting this search. Since author names were given in these fields, retrievals in keyword-in-record searches for personal names were bolstered by the addition of name elements from these fields. In Pinstripe-system searches, retrievals were also made on words in these author-name fields. Although users gave retrieved records "very useful" relevance assessments, retrieved records were not about the persons named in their original user queries.

7.5.3.3 Out-of-scope Queries for Names and Subjects

Users entered nine queries for names and subjects through personal-name subject searching capabilities. Examples of these queries are given below in terms of the elements users entered:

Last name	First name	Topic
tolstoy	peter	history
kennedy	robert	assassination
charlemagne	charlemagne	education
shakespeare	william	actors in shakespeares time
piaget	jean	psychology

The Pinstripe System submitted these queries to keyword-in-record or alphabetical searches. The Blue System tried to find both first and last name and topic elements through keyword-in-record searches. If unsuccessful, the Blue System ignored the first name element in keyword-in-record searches. As a last resort, the Blue System displayed an list of personal-name subject headings in the alphabetical neighborhood of the name elements.

In the nine searches, users searched both Blue and Pinstripe Systems. Six of the nine users began their search in the Pinstripe System. Subject searching approaches in the Pinstripe System failed to retrieve bibliographic records. Since the Blue System always responded with the alphabetical approach, it gave all nine users the opportunity to browse personal-name subject headings and retrieve bibliographic records. Three of the nine users selected listed headings. For example, the user who entered the query on Robert Kennedy selected subject headings such as "Mapplethorpe, Robert" and "Lindbergh, Charles A." The three users displayed a total of six records connected with selected name headings. Even though none of the records specifically mentioned the names they entered, users gave five "very useful" and one "not useful" ratings to the records.

7.5.3.4 Playing in Comparison Search Experiments

Users entered queries in Comparison Search Experiments that were probably not serious queries. The researchers felt users were "playing," and, thus, categorized such queries into "playing" categories. Although a total of twenty-four queries were submitted queries to searching capabilities for subjects generally, the queries were topics, names, titles, and combinations of names and topics. Examples of each are:

- "things" (topic)
- "petty" (name)
- "star trek: the lost generation" (topic)
- "Ronnie rahal the greatest baseball player ever" (combination)

Users could be quite persistent in their "playing." Two users entered six access points. In one of the two searches, the persistent user entered the access points "muslim" (twice), "pitbulls," "nursing" (twice), and "sex." The latter query matched the initial word in the subject heading "Sex discrimination in education, thus, the Blue System responded with the alphabetical approach but the user did not select a listed subject heading. The Pinstripe System responded to the latter query with a keyword-in-subdivided-heading search and displayed the same heading. The user selected this heading, displayed one record, and rated it "not useful." Even though users were "playing," the Blue and Pinstripe Systems responded with the alphabetical approach to several queries. Keyword approaches (e.g., keyword-in-record, title-keyword) were also successful in producing retrievals.

Users were probably playing when the entered twelve queries for which the system performed subject searches for personal names. Examples are queries with the following elements:

Last name	First name	Topic
Matt Van Kirk	Matt	Summer term
taraban	kimberly	love
jakeson	michael	
jakeson	michael	singing muise
kumeo	michael	ninja turtle part 3
wazir	mohammad	student

Users conducted searches in both systems except for one query which the users searched in the Blue System. Three of the eleven searches conducted in both systems began in the Pinstripe System. In twelve searches, the Blue System responded with the alphabetical approach. Even though listed personal-name headings were not on the topics specified in user queries, four users selected personal-name headings from the alphabetical list, displayed a total of nineteen bibliographic records, and rated fourteen of them "very useful," one of them "possibly useful," and the remainder "not useful." For example, the user who entered the query "taraban kimberly love" selected the heading "Taylor, Frederick Winslow," displayed seven bibliographic records, and rated all seven records "very useful."

7.5.3.5 Meaningless Input

Several types of queries constituted meaningless input: blank lines, character(s), gibberish, and letter(s). Sixty-nine subject searches and one personal-name search were characterized by blank lines. These searches were conducted in a single system. Thirty-nine searches began and ended in the Blue System; thirty-one searches began and ended in the Pinstripe System. Most likely users completed the pre-search questionnaire and walked away before entering their queries.

The character "]" was entered in a Pinstripe subject search. It retrieved no records.

A single letter or several letters were entered into fourteen subject searches. Examples of queries placed in the "letter" category are "a," "i," "j," and "s." Examples of queries placed in the "letters" category are "dD," "aa," and "aaaa." The Blue System invoked the alphabetical approach in response to several "letter" queries because the system found longer subject headings beginning with the letter entered by users. For example, the Blue System displayed an alphabetical list of subject headings beginning with the letter "a" in response to the query "a." A user chose the subject heading "APL (Programming language)," displayed over two dozen titles, gave all displayed titles "very useful" ratings, and continued selecting subject headings and cross references from the alphabetical list. The Pinstripe System responded to "letter" queries with keyword searches and the alphabetical approach. Even though their original queries made little sense, users gave displayed bibliographic records "very useful," "possibly useful," and "not useful" ratings.

Queries placed in the "gibberish" category were made up of letters, spaces, numbers, and/or characters. Nineteen subject searches and nine searches for personal names were placed in this category. Examples of gibberish are:

- jkutf[765f76b elina E\\\EEE
- fardin hajji ddJ
- vnccx6rxyrxf
- ihyfol75fg7u 6e4

Fourteen users began their searches in the Blue System and fourteen users began their searches in the Pinstripe System. Although the Pinstripe System occasionally responded with the alphabetical approach, users displayed no bibliographic records in subject searches. In three personal-name searches, both Pinstripe and Blue Systems responded with the alphabetical approach. Even though their original queries made no sense, users gave displayed bibliographic records "very useful," "possibly useful," and "not useful" ratings.

7.5.3.6 Known-item Searches

Users entered three queries through general subject searching capabilities that were probably known-item searches for titles. All three queries were similar: "schaum's" (twice) and "schaum's outline series dynamics." Since title words were included in searchable title-keyword and keyword-in-record databases, searches in Blue and Pinstripe Systems produced relevant retrievals through title-keyword and keyword-in-record searches.

Users entered a total of twenty-three queries bearing name and topic elements through personal-name subject searching capabilities that were probably requests for known items. The researchers searched the University of Michigan's Mirlyn online catalog to verify known-item requests. Although there were few queries for known items, they could be subdivided into the following specific categories: (1) out-of-scope, unverified known-item query, e.g., "jordon albert arcitecture" (sic), (2) out-of-scope, verified known-item query, e.g., "Paine Thomas Common Sense," (3) in-scope, unverified known-item query, e.g., "treddnick nick microprocessors," and (4) in-scope, verified known-item query, e.g., "kuo control systems."

Five queries produced relevant retrievals through keyword-in-record searches in Blue and Pinstripe Systems. Examples of these queries were "kuo control systems," "bastow car suspension and handling," and "juvinall robert machine design." Users retrieved relevant records because of the presence of main- and added-entry headings for personal names in the keyword-in-record database (USMARC fields 100 and 700) that the team erroneously included the searchable database of personal names.

A total of three queries were entered through subject and personal-name searching capabilities that could have been author names: "pascal," "dick," and "dijkstra." Since the user entering the query "pascal" used personal-name searching capabilities, he could have been looking for biographical information on the French scientist Blaise Pascal. In view of ASTUTE's subject coverage (computer science and technology), the user was probably searching for the computer language named Pascal. The user might not have known how to respond to the question that guided the experimental system in its choice of subject or personal-name searching capabilities because name of the programming language was the same as the name of a famous scholar.

7.5.3.7 Remaining Unusable Categories

Users entered a total of thirteen queries that the researchers placed in the "sex" terms category. Examples of "sex" queries entered through personal-name subject searching capabilities are "el-awad sex" and "Perry James sex." No queries submitted to personal-name subject searching capabilities retrieved titles. Examples of such queries entered through general subject searching capabilities were "sex (three times)," "oral sex," and "pornography." In subject searches, the query "sex" retrieved titles in the Blue System through alphabetical and keyword approaches and in the Pinstripe System through keyword approaches because the subject heading "Sex discrimination in education" was assigned to one bibliographic record. Thus, users had the opportunity to retrieve and display titles in several searches. Users gave two titles "not useful" ratings and one title a "very useful" rating in the two searches in which they displayed titles.

Users entered expletives in two searches. The lone command entered into the Pinstripe System as a query was "quit."

Three queries were computer science or engineering topics entered into the experimental online catalog through personal-name subject searching capabilities. The queries were "joint application development," "lagrange," and "dell computers." The user entering the query "lagrange" could have been looking for biographical information on the French mathematician Joseph-Louis Lagrange. However, a user entered the query "lagrangian dynamics" five minutes before the "lagrange" query. Thus, the researchers felt that the two users were the same person who wanted information on a topical subject — not on a personal name. The user entering the query with the last and first name elements "dell computers," respectively, could have responded positively to the question on persons that guided the system in its selection of personal name subject searching capabilities because company named Dell Computers sounded like it was named after a person with the last name of "Dell."

It was understandable why users responded positively to the question on persons that guided the system in its selection of personal-name subject searching capabilities. Some fine tuning to the wording of this question may be necessary to minimize the likelihood of users taking actions that invoke subject searching approaches for personal names when their queries contain proper adjectives or nouns.

Eight queries were discarded because of system errors. Three queries were discarded because a user action caused the "r" key to be mapped to the "i" key. Five were discarded because the system incorrectly coded user relevance assessments with a "0" (zero) instead of the numbers 1–3. Both errors occurred during the last week of data collection at the University of Michigan-Dearborn.

7.6 Comparison Search Experiment Administration at Earlham College

7.6.1 Introduction

The data collection period at Earlham College lasted fifteen weeks from February 23 to May 28, 1993. The experimental online catalog administered a total of 238 search experiments. The researchers manually reviewed the queries entered in these experiments to determine usable administrations for submission to data analyses. Subsections below describe the characteristics of usable and unusable search experiments. Figure 7.5 divides administrations into four components. The largest component was for usable search experiments for subject queries.

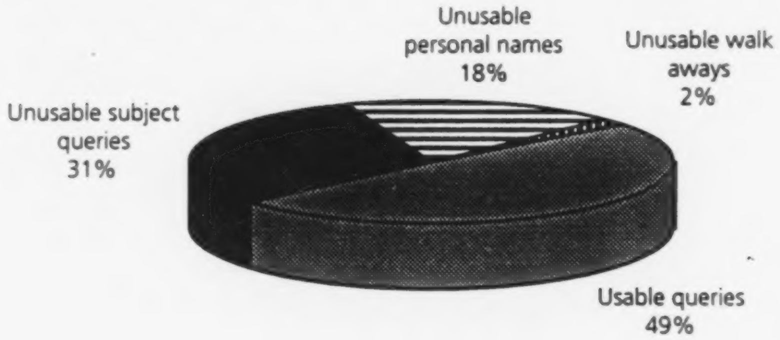


Figure 7.5. Usable and unusable Comparison Search Experiment administrations at Earlham

About half of the search administrations were usable (116 of 238 administrations). Usable search administrations include 21 administrations of the Comparison Search Experiment with library staff. A little less than a third of search administrations were unusable queries that users entered using the experimental online catalog's subject searching capabilities for subjects generally. About 18% were unusable queries entered using searching capabilities for personal names. The remainder (2%) were search administrations in which users completed one or more pre-search questions but they did not continue with their searches. These users probably walked away from the system and it eventually reset itself to the introductory screen savers. The percentages of search experiments in these four categories were comparable to percentages at UM-Dearborn. There were a greater percentage of "walk-aways" and lower percentage of unusable personal-name queries at UM-Dearborn.

Figure 7.6 depicts daily activity in terms of usable and unusable search administrations.

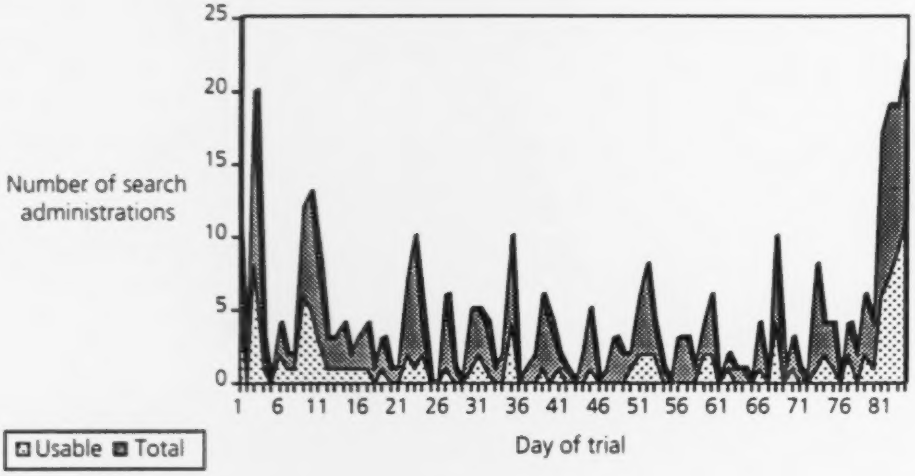


Figure 7.6. Search administrations at Earlham on a daily basis

An increase or decrease in the number of usable administrations was usually accompanied by a corresponding increase or decrease in the total number of search administrations. However, there were several days in which all search experiment administrations were deemed unusable. Toward the end of the data collection period, the number of usable administrations increased considerably. The researchers administered the Comparison Search Experiment with library staff at this time. Thus, the interviewer monitored staff searches to make sure they completed pre- and post-search questionnaires and advised them on the appropriateness of their queries in view of the experimental online catalog's subject coverage. This explains why there were many usable administrations toward the end of the data collection period.

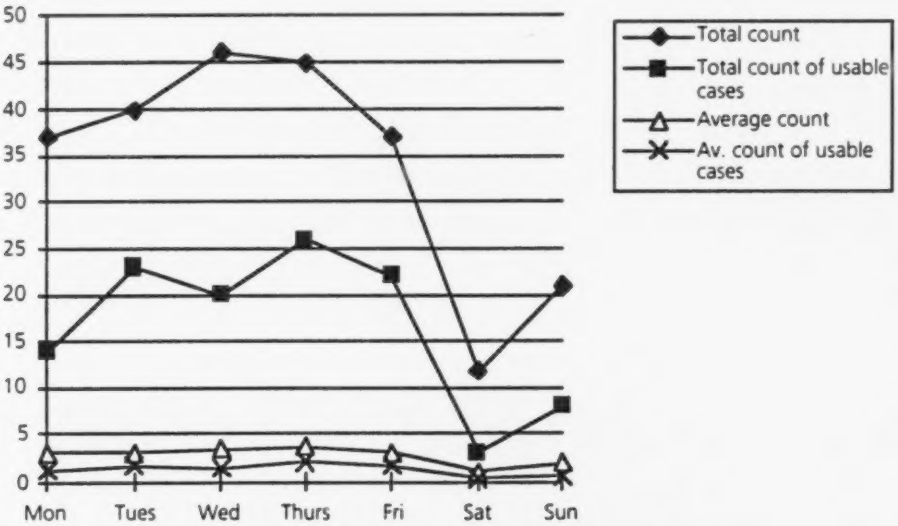


Figure 7.7. Usable search administrations at Earlham by day of the week

Figure 7.7 shows daily activity by day of the week. Searching was heavier on weekdays than on weekends.

7.6.2 Usable Administrations

Table 7.4 summarizes the characteristics of usable administrations in the Comparison Search Experiment at Earlham College.

Table 7.4. Usable Comparison Search Experiment Administrations at Earlham

Events	Completed events				
Pre-search questionnaire responses	x	x	x	x	x
Pinstripe System search	x		x	x	x
Blue System search		x	x	x	x
Post-search questionnaire responses					x
Suspect post-search questionnaire responses				x	
Number of usable search administrations (N=116)	12 (10%)	18 (16%)	20 (17%)	9 (8%)	57 (49%)

The administration of the Comparison Search Experiment at Earlham College yielded 116 usable search administrations. A large percentage (49%) of usable administrations of the Comparison Search Experiment were full administrations. Of the four partial administration categories, the largest percentage (17%) contained three complete events; the only missing event was the post-questionnaire. About 8% of administrations contained suspect post-questionnaire responses. In the data analysis, the researchers ignored suspect post-questionnaire responses and combined these administrations with search administrations in which pre-search questionnaires and searches in Blue and Pinstripe Systems had been completed. The latter administrations accounted for 17% of usable search administrations. When they were combined with the former, the two categories of usable search administrations amounted to one-quarter of usable search administrations. Remaining usable administrations had searches in only one of the two test systems (Blue or Pinstripe) and accounted for 16% and 10% of usable administrations, respectively.

Percentages of usable search administrations were similar at UM-Dearborn and Earlham. The largest percentages were for full administrations (42% at UM-D; 49% at Earlham). Administrations involving Pinstripe-System searches were more common at UM-Dearborn (14%) than at Earlham (10%). Administrations containing suspect post-search questionnaire responses were higher at UM-Dearborn (17%) than at Earlham (8%).

A total of four librarians participated in the Comparison Search Experiment at Earlham College. These librarians contributed eighteen full search administrations and two partial administrations in which they completed no post-search questionnaires. The partial administrations were database failures. Both librarian and interviewer decided it would be best to forego completing post-search questionnaires for database failures because the librarian could not draw comparisons between systems. Instead, the librarian entered new searches for different topics.

Chapters 8–12 feature the analysis of usable Comparison Search Experiment administrations.

7.6.3 Unusable Administrations

The administration of the Comparison Search Experiment at Earlham College yielded 122 unusable search administrations. Four search administrations were unusable because users completed one or more pre-search questions but they did not enter a query into the catalog. These users probably walked away from the experimental online catalog. After four minutes had elapsed, system's timeout function reset itself at the initial introductory screen savers

Table 7.5 summarizes the characteristics of the remaining 118 unusable administrations. In these administrations, users entered queries into the experimental online catalog. The queries in these unusable administrations were entered through subject searching capabilities and through personal-name subject searching capabilities. The system's choice between the two sets of capabilities was based on user responses to the question following the pre-search questionnaire that asked users whether their query involved the name of a person.

Table 7.5. Unusable Comparison Search Experiment Administrations at Earlham

	Capabilities for Subjects		**Capabilities for Personal Names**	
Query categories	**No.**	**%**	**No.**	**%**
Out of scope				
Subjects	57	76.0	2	4.7
Personal names	5	6.7	6	14.0
Names and subjects	0	0.0	29	67.4
Names, subjects, and call number	0	0.0	1	2.3
Playing				
Subjects	2	2.7	0	0.0
Meaningless input:				
Blank	7	9.3	0	0.0
Letter	1	1.3	0	0.0
Known-item search				
Title search	1	1.3	0	0.0
Author-title	0	0.0	5	11.6
Other unusable searches				
Sex term	2	2.7	0	0.0
Total	75	100.0	43	100.0

Over 80% of queries in unusable administrations were out of scope, that is, they did not match the subject coverage of the catalog's database at Earlham (i.e., American history). Since the catalog was restricted to subject searches, queries requesting author and/or titles were also considered unusable. Categories for playing, meaningless input, and other unusable searches accounted for much smaller percentages of administrations of the Comparison Search Experiment at Earlham than at UM-Dearborn. Furthermore, subcategories for gibberish, expletives, letters, command, and a few others were not needed to characterize Earlham subject queries.

7.6.3.1 Out-of-scope Subject Queries

Of the 75 queries entered through subject searching capabilities, the largest percentage (76.0%) was out of scope. Examples of these queries were "illegal abortion," "health conditions south," "koreans in japan," and "manhattan project."

The majority of out-of-scope searches were conducted in both Blue and Pinstripe Test Systems. About the same number of respondents started their searches in the Blue (19) and Pinstripe Systems (17). Users who searched in a single system averaged almost two queries per search. Users who searched both systems averaged a little over one access point per search. One searcher entered six access points: "teinmen," "steinem" (twice), "virginia woolf" (twice), and "animals."

Of the three subject searching capabilities available to the Pinstripe System for retrieving bibliographic records, the keyword-in-record search was most successful in making matches. The Blue System used five of the seven subject searching capabilities available for retrieving bibliographic records. The alphabetical approach was the most successful. The exact approach was also responsible for retrieving bibliographic records. Examples of user queries which were successful in retrieving records through the exact approach were "spiritism," "illness," "harlem renaissance" (three times), and "natural resources." Although users retrieved records for these and other queries, the researchers felt that the retrieved material did not focus entirely on their topics of interests and that users could retrieve a large quantity of material on such topics in catalog that was encyclopedic in coverage. Even though these queries retrieved bibliographic records, the researchers deemed them out of scope in a catalog limited to American history.

Twenty-one users entered out-of-scope subject queries in only one of the two experimental online catalogs. One searcher entered four access points. This searcher entered the query "woodstock." When the Pinstripe System informed the user that it was unable to find material and suggested that the user check the query's spelling, the user responded by clicking on the <Do search> three times before walking away from the experimental system. Thus, the system counted this access point four times. Only one search in single-system administrations yielded titles. Users displayed retrieved bibliographic records in only one of the two searches in each system and displayed a single title.

There were only two out-of-scope *subject queries* (1.3%) entered through *personal-name subject searching capabilities*. The queries were "woodstock" and "wood stock." One user searched both Blue and Pinstripe Systems and one user searched only the Blue System. Both users were unable to produce retrievals for these queries.

7.6.3.2 Out-of-scope Queries for Personal Names

Users entered out-of-scope personal-name queries through both general subject searching and personal-name subject searching capabilities. Three of the five personal-name queries users entered through the former were "Tanner, Henry," "marilyn monroe," and "rodin." Users searched both systems for the four of the five queries and the Blue System only for the "rodin" query. Neither the Blue nor the Pinstripe System yielded retrievals for these queries.

Users entered six out-of-scope personal-name queries through personal-name subject searching capabilities in the experimental online catalog at Earlham. Users searched all six queries in the Blue Test System. These six out-of-scope personal-name queries were:

User query	Selected personal-name heading
singer	
chopin kate	
poe edgar	
einstein albert	
resnick	Woodring, Harry Hines.
stone harlan	Stoner, Anna Louisa Wellington.

As a last resort, the Blue System always responded to personal-name queries with the alphabetical approach. In these six searches, the system responded with this approach which showed users a list of personal-name headings in the alphabetical neighborhood of entered name element(s). In two of six alphabetical searches, users selected a listed heading and displayed one bibliographic record connected with the heading. Although the original user queries and selected personal-name headings seemed to have nothing in common, users gave "very useful" and "possibly useful" ratings to retrieved titles.

7.6.3.3 Out-of-scope Queries for Names and Subjects

Users entered twenty-nine queries for names and subjects through personal-name subject searching capabilities. Examples of these queries are given below in terms of the elements users entered:

Last name	First name	Topic
debs	eugene	labor
houdini	harry	spirit mediums
charlemagne	charlemagne	education
socrates		philosiphy (sic)

jones	jim	biography

The Pinstripe System submitted these queries to keyword-in-record searches. The Blue System tried to find both first and last name and topic elements through keyword-in-record searches. If unsuccessful, the Blue System ignored the first name element in keyword-in-record searches. As a last resort, the Blue System displayed an list of personal-name subject headings in the alphabetical neighborhood of the name elements.

In the twenty-nine searches, three users searched a single system once — the Blue System. Of the users searching both systems, a few more users started in the Pinstripe System (14) than in the Blue System (12). Subject searching approaches in the Pinstripe System failed to retrieve bibliographic records. Since the Blue System always responded with the alphabetical approach, it gave users the opportunity to browse personal-name subject headings and retrieve titles. Nine users selected listed headings. Examples of original user queries and selected personal-name subject headings were:

User query	Selected personal-name heading
stanton elizabeth suffrage	Steele, Eliza R.
welter barbara women's history	Welles, Gideon
nash diane civil rights movement	Nasser, Gamal Abdel
Baker Ella Student Non-violent coordinating committee	Baker, Newton Diehl
Jackson LaToya The life of LaToya Jackson	Jackson, Andrew; Jackson, Rachel

Eight of nine users who selected listed personal-name headings from alphabetical lists also displayed titles connected with them. Of the twelve titles displayed, users rated almost all "not useful." Although they rated two records "very useful," it was reassuring that users were taking relevance assessments seriously in these instances.

One user query was a combination of call number, name, and subject (or title) elements. The query was "pr 935 s2.6" (last name element), "christy " (first name element), and "cooking or dogs" (topic element). Although the Blue System displayed personal-name headings beginning with the letter "p," the user did not select one.

7.6.3.4 Remaining Subject Queries

The thirteen remaining subject queries were distributed into six query categories. Examples of queries entered in searches placed in the "playing" were "lizzie's big butt," "heidi," and "poop." Both Blue and Pinstripe Systems were unsuccessful in producing retrievals for these queries.

Seven searches were characterized by blank lines. These searches were conducted in a single system with five users searching the Blue System and two users searching the Pinstripe System. In these search administrations, users probably completed the pre-search questionnaire and walked away before entering their queries.

In the one search bearing a single letter ("s"), the keyword-in-subdivided-heading search of the Pinstripe System and the alphabetical approach of the Blue System were successful in displaying subject headings. The user, however, did not select listed heading(s).

The two queries placed in the "sex" category failed to retrieve titles in the Blue or Pinstripe Test Systems.

A single user entered a query for a known item using general subject searching capabilities. The user entered the query "the age of jackson." After submitting it to exact and alphabetical approaches, the Blue System retrieved nine bibliographic records through a title-keyword search because title words make up the searchable database for this search. Of the nine records, the user gave one record a "very useful" rating and the other eight records "possibly useful" and "not useful" ratings.

7.6.3.5 Remaining Personal-name Queries

Users entered a total of five queries for known items using personal-name searching capabilities. Although two queries were entered on different days, they could have been entered by the same user because they asked for the same materials. The queries were "schlesinger aurthur urban history." A similar query "schlesinger arthur the age of jackson" was entered about the same time as the other two queries but not immediately before or after them. The query "parker sarah comanche indians" could have been an author-title search but the researchers could not verify this item in searches of operational online catalogs. Through the Blue System's alphabetical approach, users retrieved titles for two of the five queries but they gave them "not useful" ratings.

7.7 Reliability of Relevance Assessments

This was not the first study of user queries to discover that users entered terms that were not serious subject queries. Several studies (Drabenstott and Vizine-Goetz, 1994; Peters 1989; Henty 1986; Walter 1987; Hunter 1991; Dickson 1984; Lester 1991) characterized such queries as "sex terms," "malicious entries," "gibberish," "garbage," "wrong file," etc. Although it seemed unlikely that such queries would produce retrievals, there were several situations in which the Blue and Pinstripe Systems gave users opportunities to display subject headings and bibliographic records. For example, the Pinstripe System gave users the opportunity to select subject headings and display bibliographic records connected to them because it randomly selected the alphabetical approach in response to user queries regardless of the extent to which they matched listed headings. In personal-name searches, the Blue System responded with the alphabetical approach to name-query elements as a last resort. Keyword searches in both systems were sometimes successful in retrieving records for single letters or sex terms.

The discouraging factor with these retrievals was the preponderance of "very useful" ratings users gave to displayed bibliographic records. Let us illustrate this phenomenon in examples of two searches. In the first example, a user at UM-Dearborn entered the single letter query "j" into the Pinstripe System. The system displayed two titles retrieved in a keyword-in-record search. The user gave one title a "very useful" rating and the other title a "possibly useful" rating even though the two records were on totally different subjects, viz. CAD and jet propulsion systems. In the second example, an Earlham user entered the surname "resnick" into the experimental online catalog. The Blue System responded with the alphabetical approach. The user selected the personal-name heading "Woodring, Harry Hines," displayed the one title connected with it, and gave it a "useful" rating even though it did not mention a person named resnick. (The retrieved title was a biography written by Keith D. McFarland of Harry H. Woodring, Secretary of War under Franklin D. Roosevelt.)

These examples and the relevance assessments made by many other users in searches for unusable queries made us skeptical about the reliability of relevance assessments and post-search questionnaire responses. We examined the method we used to gather such data to determine how it could be changed in future administrations of online questionnaires.

Figure 6.11 shows the Pinstripe System's display of a title retrieved in a search for "civil rights" bearing a pop-up window that gives users three categories for rating usefulness: (1) "very useful," (2) "possibly useful," and (3) "not useful." The user could click on one of these three categories using the mouse or use a combination of arrow keys and <Enter> key to select a category. The first-listed category (i.e., "very useful") was given in white type and set in a medium-orange color against a light-blue background. The other categories were not highlighted but they were given in light yellow type and set in a navy blue background. The simplest action for the user to take was to hit the <Enter> key. This made the system choose the first-listed category, i.e., "very useful" and record it to level 3 transaction log records for relevance assessments. Reviewing relevance assessments for unusable queries, the researchers felt that many users performed this simple action in response to relevance assessment pop-up boxes. This action could have been more prevalent in searches for unusable queries that were meaningless, were for play, sex terms, or expletives, because users were not serious about their searches.

If the ASTUTE project team could redesign the relevance assessment pop-up window, they would make the selection process a much more deliberate step on the part of the user than hitting the <Enter> key. There are several possibilities and combinations of possibilities for redesigning this window. If the first-listed category were to remain highlighted, it should be changed to a category such as "not useful" or "don't know." Thus, a user's selection of the first-listed category by the simple action of hitting the <Enter> key would not "stack the deck" in favor of "useful" relevance assessments but it would allow for a conservative estimate of useful titles. Another approach would involve requiring users to type in a number or letter corresponding to the category of their choice instead of using arrow keys to select listed categories. This would make their selection more deliberate than repeatedly hitting the <Enter> key. The researchers also felt that users would be more conscientious about selecting categories that reflected the relevance of displayed items if they knew that the system would use such assessments to find additional titles. Thus, online catalogs that feature relevance assessment capabilities should inform users that their assessments will be used for finding additional titles to ensure that users give serious consideration to their relevance assessments.

7.8 Reliability of Post-search Questionnaire Responses

The experimental online catalog always concluded search administrations with post-search questionnaires. Although many users whose search administrations were deemed unusable failed to complete post-search questionnaires, about the same number completed them or most of them. Many users who entered out-of-scope queries were serious about their questionnaire responses in that they chose categories that documented the system's inability to produce retrievals for their queries. Other users such as those who were not serious about their queries (e.g., queries in "letter," "sex term," "gibberish" categories) were not serious about their responses to the questionnaire and chose responses that did not make sense in view of their search results. For example, there were searches in which neither system retrieved titles but users chose answers to post-search questions that implied that one or both systems retrieved titles.

Another problem with post-search questionnaire responses was similar to the problem with relevance assessments. In unusable *and* usable search administrations, users repeatedly selected the first-listed response category in post-search questionnaires. A manual review of searches and relevance assessments made the researchers skeptical about the reliability of these post-search questionnaire responses. We examined the method used to gather such data to determine how it could be changed in future administrations of online questionnaires.

The experimental online catalog displayed each question and response categories separate from other questions and response categories. The first-listed category was given in a bright yellow type and set in a medium-orange color against a dark blue background. The other categories were given in a light yellow type against a dark blue background. The simplest action for the user to take was to hit the <Enter> key. This would make the system choose the first-listed category and save it to level 1 transaction log records.

In the analysis of usable search administrations (see chapters 8–12), the researchers disregarded post-search questionnaires in which users selected most or all first-listed response categories.

If the ASTUTE project team could redesign the post-search questionnaire, they would make the selection process a much more deliberate step on the part of the user than hitting the <Enter> key. There are several possibilities and combinations of possibilities for redesigning the selection process. If the first-listed category were to remain highlighted, it could be changed to a category such as "don't know" or "don't care." First-listed categories could also be different from question to question. When users selected one or more such categories, the truthfulness of their responses to other post-search questions could be in doubt. Another approach would involve requiring users to type in a number or letter corresponding to the category of their choice. This would make their selection more deliberate than repeatedly hitting the <Enter> key.

7.9 Chapter Summary

This chapter describes the manual review of search administrations that the researchers performed to determine whether patron search topics were in scope and whether their responses to pre- and post-search questionnaires were valid.

The three installations of the experimental online catalog lasted for five- to fifteen-week periods at the two participating libraries. Data collected in the first of two data-collection periods at UM-D were discarded because of a system error that corrupted data for over two-fifths of the search administrations (see section 7.4). The system error was introduced by the unexpected user action of repeatedly turning the microcomputer on and off. The programming team made changes to the experimental online catalog to ensure that the data would not become corrupted in later system installations.

In the second, five-week data-collection period at UM-D, the experimental online catalog administered a total of 826 search experiments (see section 7.5). In the fifteen-week data-collection period at Earlham, the experimental online catalog administered a total of 238 search experiments (see section 7.6).

The manual review of searches revealed that about 50% of search administrations at UM-D and Earlham were usable (see sections 7.5.1 and 7.6.1); however, the manual review of post-search questionnaires and questionnaire responses revealed that large percentages (58% at UM-D, 51% at Earlham) of search experiments were less-than-full administrations (see sections 7.5.2 and 7.5.3). That is, users completed a search in one system, both systems, and/or failed to answer the post-search questionnaire. Since monitors were not present to encourage patrons to complete the entire Comparison Search Experiment, the researchers expected patrons to leave before they finished the full administration of the Comparison Search Experiment. The manual review process was essential because it identified full and less-than-full search administrations which were, on occasion, handled differently in the statistical analysis of search results (see chapter 9). The failure analysis described in chapters 10–12 focuses on both full and less-than-full usable search administrations.

Unusable administrations were divided into categories for queries that were out of scope, for play, meaningless input, known-item searches, and other unusable searches (e.g., sex terms, expletives) (see sections 7.5.3 and 7.6.3). Discussions of unusable search administrations revealed that users selected subject headings, displayed titles, and rated some titles "very useful" even though they were not about the topics users originally entered in their searches. Discouraged with the positive relevance assessments users gave to displayed titles in unusable search administrations, the researchers made suggestions for redesigning the procedures for collecting relevance assessments in future experimental and operational systems (see section 7.7). They also made suggestions for redesigning system-administered post-search questionnaires to minimize the collection of unreliable data in future experimental systems (see section 7.8).

References

Dickson, Jean. 1984. "An analysis of user errors in searching an online catalog." Cataloging & Classification Quarterly 4, 3 (Spring): 19–38.

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Henty, Margaret. 1986. "The users at the online catalogue: A record of unsuccessful keyword searches." *LASIE* 17, 2 (September/October): 47–52.

Hunter, Rhonda N. 1991. "Successes and failures of patrons searching the online catalog at a large academic library: A transaction log analysis." *RQ* 30, 3 (Spring): 395–402.

Kaske, Neal K. 1988. "A comparative study of subject searching in an OPAC among branch libraries of a university library system." *Information Technology and Libraries* 7, 4 (December): 359–72.

Lester, Marilyn Ann. 1989. *Coincidence of user vocabulary and Library of Congress Subject Headings: Experiments to improve subject access in academic library online catalogs*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Peters, Thomas A. 1989. "When smart people fail: An analysis of the transaction log of an online public access catalog." *Journal of Academic Librarianship* 15, 5 (November): 267–73.

Walter, Dennis R. 1987. "The user at the online catalogue: A record of unsuccessful keyword searches — another case study." *LASIE* 18, 3 (November/December): 74–81.

8 Characteristics of System Users and Searches

8.1 Introduction

Chapter 8 introduces readers to the characteristics of system users and searches. Characteristics of system users are gleaned from pre-search questionnaire responses. Search characteristics come from an analysis of "usable" search administrations. The analysis reveals interesting findings about the time spent searching, length of user queries, and matches of user queries with the experimental system's controlled vocabulary. Search characteristics include an analysis of titles displayed in the Blue and/or Pinstripe System. The analysis is necessary to identify titles retrieved more than once by one or both systems and such titles given conflicting relevance assessments. This chapter concludes with decisions that the ASTUTE project team made to resolve conflicts in preparation for the statistical analysis in the chapter that follows.

8.2 Experimental System Users

All respondents whose questionnaires were deemed "usable" answered the three questions in the pre-search questionnaire. The first question in the pre-search questionnaire asked respondents how many times they had used the experimental ASTUTE online catalog. Figure 8.1 summarizes respondent answers at UM-Dearborn and Earlham College to the first pre-search question, viz. "How many times have you tested this system?" Pre-search questionnaire responses came from 412 users at UM-D and 116 users at Earlham.

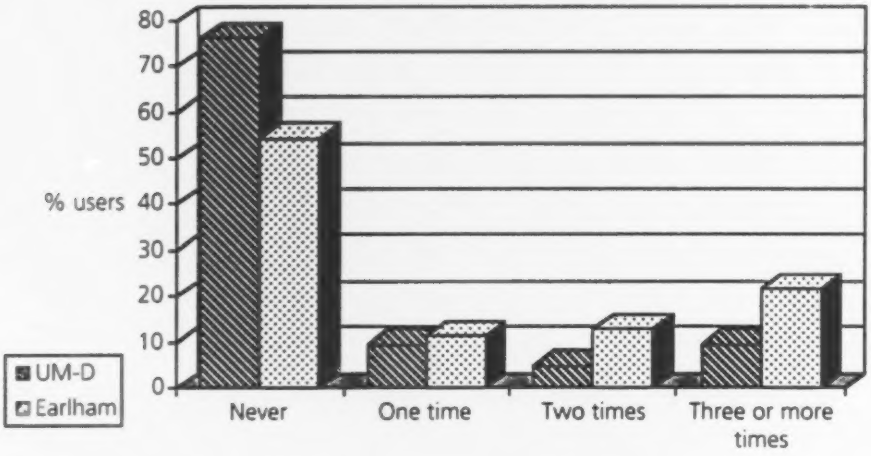


Figure 8.1 Previous use of ASTUTE

About three-quarters of UM-D users were searching ASTUTE for the first time. A little less than half of Earlham users were repeat users.

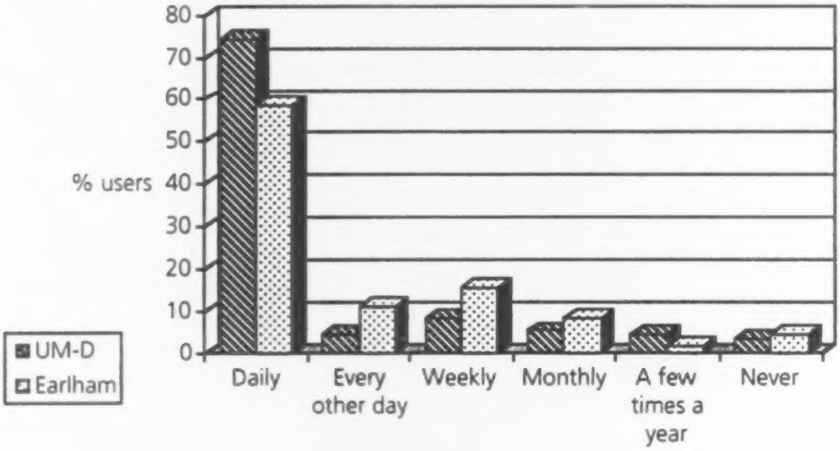


Figure 8.2. Other computer use

Figure 8.2 summarizes how respondents completed the statement "I use a computer system other than ASTUTE" which was the second pre-search question.

A large percentage (74.0%) of systems users at UM-D used computers on a daily basis. Such a high percentage was anticipated because ASTUTE's database contained materials on computer science and engineering. Students who studied these fields were likely to be frequent computer users. System users at Earlham were also frequent computer users. Almost seventy percent of Earlham respondents used computers daily or every other day.

Table 8.1 gives the major field of study of experimental system users at UM-D and Earlham.

Table 8.1. Major Field of Study

Major field of study	UM-D (n=412)	Earlham (n=116)
Computer science	29.1	12.9
Engineering	46.1	0.0
Business	3.6	1.7
Education	0.5	2.6
History	0.5	25.0
Mathematics	1.2	0.9
Other humanities fields	1.9	12.9
Other social science fields	1.0	11.2
Other science fields	4.4	4.3
None of the above	11.7	28.5
Total	100.0	100.0

Large percentages of UM-D respondents were computer science or engineering majors (29.1% and 46.1%, respectively). Thus, the majority of UM-D respondents were performing searches for topics in their major field of study. This was not true of respondents at Earlham College where ASTUTE's database focused on American history. Only one-quarter of Earlham respondents were history majors; consequently, the majority of Earlham respondents were not performing searches for topics in their major field of study.

A little under one-quarter of Earlham respondents were majoring in fields in the humanities or social sciences (12.9% and 11.2%, respectively). Although we have no details on the specific fields of respondents who selected these categories, such respondents might be familiar with history from coursework in cognate subjects. The large percentage of computer science majors at Earlham is suspect. This major was the first in a list of ten categories. Respondents could have been in a hurry to complete the pre-search questionnaire and hit the <Return> key (which selected the first-listed category, i.e., computer science) rather than moving the arrow keys down to the appropriate category and hitting the <Return> key. (See section 7.8 for a discussion of suspect questionnaire responses.)

8.3 Time Spent Searching

Tables 7.2 and 7.4 described the events that respondents completed in usable search administrations. Respondents could have completed as few as two events, i.e., pre-search questionnaire and one search in the Blue- or Pinstripe-test System, or as many as four events, i.e., pre-search questionnaire, one search in the Blue-test System, one search in the Pinstripe-test System, and post-search questionnaire. Table 8.2 summarizes the average time it took respondents to complete each type of search administration.

Table 8.2. Time Spent Per Search Administration Type

Events	**Completed events**				
Pre-search questionnaire responses	x	x	x	x	x
Pinstripe System search	x		x	x	x
Blue System search		x	x	x	x
Post-search questionnaire responses					x
Suspect post-search questionnaire responses				x	
Time spent searching					
UM-D	4.3	4.6	4.9	4.6	6.4
Earlham	4.3	4.2	4.4	3.0	6.4

Respondents who searched a single system or who searched two systems but failed to complete the post-search questionnaire spent between four and five minutes searching. Respondents who searched both Blue and Pinstripe Systems and completed a post-search questionnaire spent the longest time searching — 6.4 minutes.

Search administrations with suspect responses to post-search questionnaires consisted of administrations with default responses that users could select by repeatedly pressing the <Enter> key or by repeatedly entering the same number. At Earlham, such search administrations were quite short — three minutes. At UM-D, the time spent searching in such administrations was comparable to the times spent searching a single system.

Figures 8.3 and 8.4 show the average length of time users took to complete each event in search administrations: (1) searches in the Blue and/or Pinstripe Systems (designated "Search"), (2) three pre-search questions (designated "Pre-search"), and

(3) eleven post-search questions (designated "Post-search") at UM-D and Earlham, respectively. Average search times summarized in figures 8.3 and 8.4 were calculated from 412 searches at UM-D and 116 searches at Earlham.

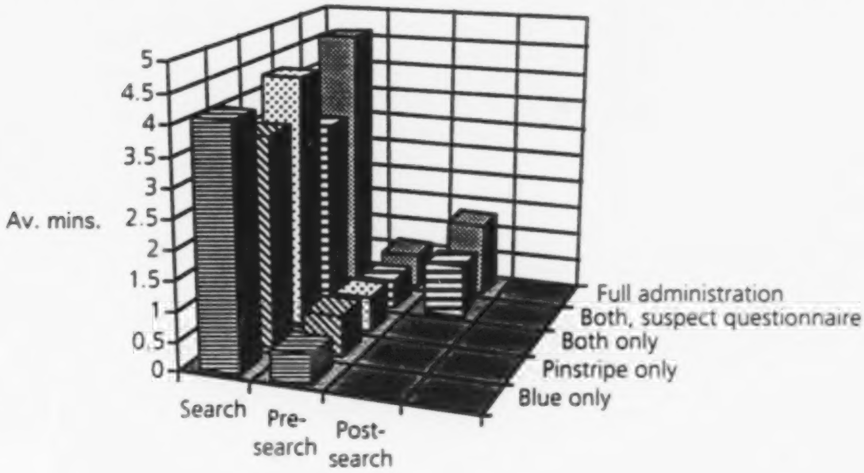


Figure 8.3. Time spent searching per event at UM-D

Searchers spent less than a minute completing the pre-search questionnaire. Time spent searching was quite low for search administrations bearing suspect post-search questionnaires — a little more than three minutes at UM-D and two minutes at Earlham. Time spent searching was the longest for full search administrations — over four minutes at both UM-D and Earlham. Completing post-search questionnaires took less than a minute in search administrations with suspect questionnaire responses and took a minute and a half in full search administrations.

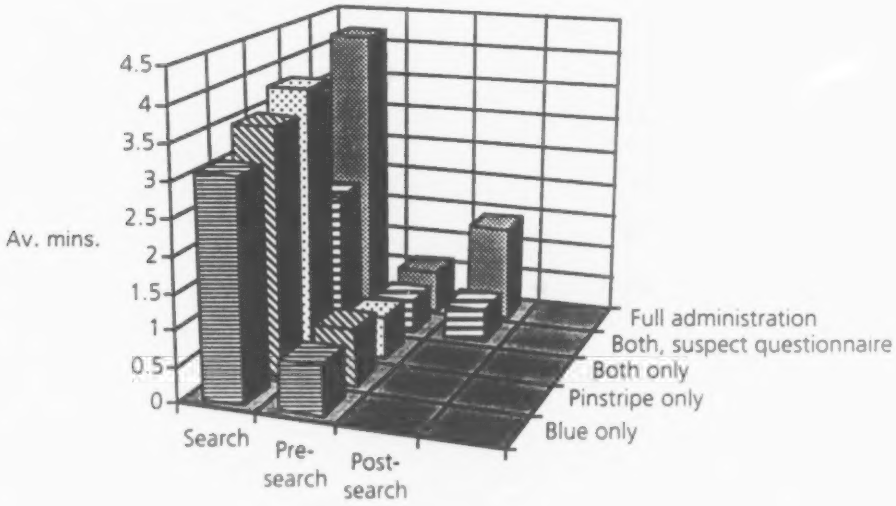


Figure 8.4. Time spent searching per event at Earlham

Since end users conducted searches in two systems, it was entirely possible that they would tire of searching the experimental online catalogs and truncate their searches in the second catalog. To determine whether this was true, we reviewed the time spent searching the two systems. Table 8.3 gives the time spent searching the Blue and Pinstripe Systems according to whether they were the first or second system searched. (Times given in Table 8.3 do *not* include search administrations in which users searched only one of the two test systems.)

Table 8.3. Time Spent Searching Each System

	First system searched		Second system searched		Total by library	
Library	**Blue**	**Pinstripe**	**Blue**	**Pinstripe**	**Blue**	**Pinstripe**
UM-D	2.9	2.1	1.4	1.9	2.1	2.0
Earlham	2.8	1.2	1.9	2.5	2.3	1.9
Total by system	2.8	1.9	1.5	2.1	2.2	2.0

When UM-D users began their search in the Blue System, they spent 2.9 seconds searching this system and 1.9 seconds searching the Pinstripe System. When UM-D

BEST COPY AVAILABLE

users began their search in the Pinstripe System, they spent 2.1 seconds searching this system and 1.4 seconds searching the Blue System. Thus, at UM-D, users spent more time searching the first system than the second system. When Earlham users began their search in the Blue System, they spent 2.8 seconds searching this system and 2.5 seconds searching the Pinstripe System. When Earlham users began their search in the Pinstripe System, they spent 1.2 seconds searching this system and 1.9 seconds searching the Blue System. The latter users actually spent more time searching the second system than the first system.

Overall, users spent about two minutes searching each of the two experimental systems. Searches in the Blue System ranged from less than a half minute to about twenty minutes. One UM-D user entered the query "radiator design" into the Blue System. When the Blue System failed to make a match, the user entered the query "heat transfer." The Blue System responded with the exact approach to this cross-reference. The user spent the next twenty minutes browsing subdivided forms of the matched heading, browsing narrower terms, displaying titles assigned to the matched heading and to narrower terms, displaying over fifty titles, and giving varied relevance assessments to retrieved titles. Although the keyword-in-record approach in the Pinstripe System retrieved about ninety titles for the user query "heat transfer," this user only spent a little over two minutes reviewing the results in the Pinstripe System.

A few searches in the Pinstripe System lasted as long as thirty minutes. One UM-D user began searching in the Blue System for "composites" and spent two and a half minutes browsing the results of the alphabetical approach. When this user terminated his search of the Blue System, ASTUTE responded with the results of an alphabetical search in the Pinstripe System. The user browsed the alphabetical list, selected terms of interest, browsed broader and narrower terms connected with selected term, displayed over eighty titles, and varied his relevance assessments of displayed titles. His search of the Pinstripe System lasted almost twenty-nine minutes.

8.4 User Queries

ASTUTE was designed to handle one query per search experiment administration. That is, it prompted users to enter a query, gave search results in the first system, responded to a user action to start over or terminate the search by switching to the second system and automatically giving search results in the second system, responded to a user action to start over or terminate the search by initiating the post-search questionnaire, and terminated the search at the completion of the questionnaire. Thus, the average number of queries per search was expected to be low in comparison to studies of operational online catalogs (Markey 1984, 66; Drabenstott and Vizine-Goetz 1994, 169, 186).

When ASTUTE failed to make matches of one or more words in user queries, the system prompted users to check and correct spelling. Some users responded by entering entirely different queries. Table 8.4 summarizes characteristics of user queries.

Table 8.4. Characteristics of User Queries

	UM-D		Earlham	
Characteristics	**Blue**	**Pinstripe**	**Blue**	**Pinstripe**
No. of queries (av.)	1.3	1.3	1.1	1.1
No. of queries (range)	1–8	1–6	1–8	1–6
No. of query words (av.)	1.8	1.6	2.9	3.0
No. of query words (range)	1–7	1–6	1–12	1–11
No. of words in queries that make matches (av.)	1.5	1.5	2.1	2.4
No. of words in queries that fail to make matches (av.)	2.0	1.9	3.2	4.1

At UM-D and Earlham, the average number of queries per search was about the same, about 1.1 or 1.3 queries per search. This low average was expected because of ASTUTE's unique design.

At UM-D, the average number of query words was almost two words. At Earlham, the average number of words in queries was almost three words. The average number of words in queries entered by UM-D respondents was typical for queries for topical subjects and geographical names that were likely to produce retrievals in online catalogs (Drabenstott and Vizine-Goetz 1994, 169, 186). The average number of words in queries entered by Earlham respondents was typical for queries for queries that were not likely to produce retrievals, e.g., queries for personal names, for combinations of names and topics (Drabenstott and Vizine-Goetz 1994, 200, 228, 234).

Queries ranged from one to twelve words. Examples of long queries are:

- digital design with verilog hdl
- fundamental of data structure in c
- king (surname); coretta (given name); scott (middle name); learn about what happened in "1927 "

- king (surname); martin (given name); luther (middle name); about the bycotts they had way back in Mississippi (topic)
- monitoring performance of telephone operators
- women in the westward movement

The last two rows in Table 8.4 give the average number of words in queries that made matches and failed to make matches in the Blue or Pinstripe Systems. When queries made matches in the Blue or Pinstripe Systems, the systems responded with the results of a particular subject searching approach. That is, the systems retrieved titles or subject headings bearing words in user-entered queries. At both UM-D and Earlham, queries that failed to make matches were longer than queries that made matches.

We compared the average length of queries that made matches to the average length of queries that failed to make matches to determine if the difference between the lengths was significant. The results of t-tests examining query length and matches are given in Table 8.5.

Table 8.5. T-tests Comparing Query Lengths

Match type	Number	Mean length	Standard deviation
T-test results for Blue System at UM-D			
No match	172	1.9709	1.0284
Successful match	80	1.5125	0.6363
T = -3.6713	Degrees of freedom = 250		Probability = 0.0003
T-test results for Pinstripe System at UM-D			
No match	121	1.8595	0.9601
Successful match	164	1.4512	0.6296
T = -4.3297	Degrees of freedom = 283		Probability = 0.0000
T-test for results for Blue System at Earlham			
No match	43	3.2326	2.3077
Successful match	16	2.1250	1.2583
T = -1.8153	Degrees of freedom = 57		Probability = 0.0747
T-test for results for Pinstripe System at Earlham			
No match	25	4.1200	2.2420
Successful match	44	2.4091	1.4518
T = -3.8469	Degrees of freedom = 67		Probability = 0.0003

All t-tests gave significant results except for the t-test on mean lengths of queries in the Blue System at Earlham. Although the difference in mean lengths was greater than differences in the mean lengths for queries entered into the experimental online catalogs at UM-D (i.e., 1.1076 words vs. 0.4584 or 0.4083 words), the t-test failed to achieve significance. Its failure to achieve significance could be due to the low sample size for Earlham queries in the Blue System. Generally speaking, the length of queries had a significant effect on the ability of the experimental online catalog to make matches of its subject vocabulary. The longer queries were, the less likely the experimental systems would make a match.

8.5 Retrieved and Displayed Titles

8.5.1 Opportunities to Retrieve Titles

The Blue and Pinstripe Systems responded to user-entered queries that matched indexed subject terms with controlled vocabulary approaches (e.g., exact, alphabetical, keyword-in-heading) or free-text approaches (e.g., title-keyword, keyword-in-record). The former approaches provided users with intermediary browsing lists of subject headings from which users selected headings of interest. The Blue System's selection of a particular subject searching approach was governed by search trees that required queries to make various matches of controlled vocabulary or free-text terms. The Pinstripe System selected approaches randomly.

Search administrations in the Blue or Pinstripe System that contained a "1" in the L2_Results field of transaction log records gave users the opportunity to display bibliographic records. Search administrations in the Blue or Pinstripe System that contained a "0" in the L2_Results field did not give users the opportunity to display bibliographic records (Table F3 in Appendix F). The latter searches were roughly equivalent to searches in which zero titles were retrieved.

ASTUTE always wrote a "1" to the L2_Results field when the Pinstripe System chose the alphabetical approach at random in response to user queries regardless of the closeness of user queries to subject headings in the alphabetical list. For example, the Pinstripe System responded with the alphabetical approach to the user query "treatment of women slaves during the civil war." Examples of listed subject headings were "Treaties," "Treaty of Ghent," "Trenton, Battle of, 1776," and "Trial by jury;" these subject headings had nothing to do with the user's topic of interest. Thus, the existence of a "1" in the L2_Results field did not necessarily mean that the Pinstripe-system's response (subject headings or titles) was useful to users.

In Table 8.6, Blue and Pinstripe searches are categorized according to search administration type. Number columns give the number of search administrations per search administration type. Percentage columns give the percentage of searches which gave users the opportunity to retrieve and display titles (that is, the percentage of searches in which the system made a match, and, thus, wrote a "1" to the L2_Results field of transaction log records). Table 8.6 summarizes the percentages of Blue and Pinstripe searches that gave users the opportunity to pursue the results of subject searching approaches.

Table 8.6. "Opportunities" by Search Administration Type

	% UM-D				% Earlham			
	Blue		Pinstripe		Blue		Pinstripe	
Search administration type	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Search in two systems, no or suspect questionnaire	126	77.8	126	56.3	29	75.9	29	62.1
Search in two systems, questionnaire	171	78.9	171	73.7	57	84.2	57	54.9
Total	353	79.3	356	73.9	104	83.7	98	62.2

Key:

No. = number of searches in this search administration type

% = percentage of searches with opportunities

Table 8.6 omits results for searches in a single system. Percentages of match for search administrations in which users conducted a search in a single system cannot be compared to one another because the searches in the Blue and Pinstripe Systems were for different topics.

Percentages of opportunities were lower in the Pinstripe System than in the Blue System. The percentages for the Pinstripe System (at both institutions) would have been even lower if we had subtracted the number of Pinstripe-system searches in which the alphabetical browsing list of subject headings produced by the alphabetical approach contained subject headings that did not match the initial word(s) in longer subject headings.

Of the analyses described so far, the analysis summarized in Table 8.6 provides the first comparison of system performance. The analysis demonstrated that the Blue System was likely to provide more opportunities for guiding users to useful information on their topics of interest than the Pinstripe System provided.

8.5.2 Retrieving Titles

Table 8.7 summarizes the number of titles retrieved in Blue and Pinstripe searches and categorizes titles by the particular search administration type.

Table 8.7. Retrievals by Search Administration Type

	% UM-D				% Earlham			
	Blue		Pinstripe		Blue		Pinstripe	
Search administration type	**No. searches**	**No. titles**	**No. searches**	**No. titles**	**No. searches**	**No. titles**	**No. searches**	**No. titles**
Search in one system only	56	983	59	3,124	18	242	12	112
Search in two systems, no or suspect questionnaire	126	1,969	126	5,222	29	220	29	1,076
Search in two systems, questionnaire	171	2,967	171	7,927	57	694	57	736
Total	353	5,919	356	16,273	104	1,156	98	1,924

Pinstripe-system users retrieved many more titles per search than Blue-system users. This was especially evident for UM-D users who retrieved over two times the number of titles in the Pinstripe System than in the Blue System. High-posted searches in the Pinstripe System were likely to come from the keyword-in-record search which was one of the three search approaches implemented in this system. Keyword-in-record searches made retrievals based on matches of words in all subject-rich fields (see section 5.8.3).

8.5.3 Displaying Retrieved Titles

Providing users the opportunity to display titles did not guarantee that users would choose seemingly relevant subject headings from a list that would take them one step closer to browsing titles connected with the subject heading. Furthermore, it did not guarantee that users would browse more than the first title displayed in free-text approaches.

For example, an Earlham user entered the personal name query "jefferson" (last name), "thomas" (first name), and "inventions" (topic) into the Blue System. Exhausting controlled vocabulary and free-text approaches for name and topic elements, the Blue System displayed an alphabetical list of subject headings for personal names that contained the name "Jefferson, Thomas, 1743–1826." The user selected this name but she did not display one of the several dozen titles bearing this name.

Failing to take advantage of the opportunity to display seemingly relevant bibliographic records was not an uncommon phenomenon. It occurred in both Blue and Pinstripe searches and amongst users at both institutions. In the failure analysis of experimental system searches (chapters 10–12), this problem was called "perseverance" and was studied in depth to determine why it happened and what system design changes would minimize it (see sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1).

Table 8.8 summarizes the number and percentage of titles users retrieved and displayed in Blue- and Pinstripe-system searches.

Table 8.8. Retrieved and Displayed Titles

	% UM-D				**% Earlham**			
	Blue		**Pinstripe**		**Blue**		**Pinstripe**	
Search administration type	No. titles retrieved	% titles displayed	No. titles retrieved	% titles displayed	No. titles retrieved	% titles displayed	No. titles retrieved	% titles displayed
Search in one system only	983	29.9	3,124	10.2	242	35.1	112	45.5
Search in two systems, no or suspect questionnaire	1,969	26.3	5,222	13.6	220	40.9	1,076	9.4
Search in two systems, questionnaire	2,967	36.6	7,927	12.9	694	41.9	736	21.2
Total	5,919	32.1	16,273	12.6	1,156	40.3	1,924	16.0

Experimental-system users retrieved fewer titles in the Blue System than in the Pinstripe System but they displayed a larger percentage of titles per search in the Blue System than in the Pinstripe System. Blue-system users at UM-D displayed nearly one-third of the titles they retrieved. Pinstripe System users at UM-D displayed about one-eighth of the titles they retrieved. At Earlham, Blue-system users displayed about 40% of the titles they retrieved. Overall, Pinstripe-system users at Earlham displayed only 16% of the titles they retrieved.

8.5.4 Duplicate Titles in Retrievals

Every time the Blue and Pinstripe Systems displayed a title, they asked searchers to rate the usefulness of the title and gave them three usefulness responses to select: (1)

"very useful," (2) "somewhat useful," and (3) "not useful." When users displayed titles more than once, they occasionally gave different relevance assessments. Statistics on displayed titles included titles that users displayed more than once (Tables 8.7 and 8.8). Searchers had many opportunities to display the same title time again and again. Examples of such opportunities are:

- Repeating the display of previously displayed title(s) by selecting <Previous title> or <Next title> options
- Retrieving and displaying the same title(s) in subsequent Blue-system searches invoked by user selection of the <Expand search> option.
- Retrieving and displaying the same title(s) in a Pinstripe-system search following a Blue-system search or in a Blue-system search following a Pinstripe-system search.

At UM-D, Blue- and Pinstripe-system users displayed a total of 470 titles more than once. Of this total, 82.8% of titles were displayed twice across systems, 11.5% were displayed three times across systems, and 5.7% were displayed more than three times across systems. At Earlham, Blue- and Pinstripe-system users displayed a total of 183 titles more than once. Of this total, 80.1% of titles were displayed twice across systems, 12.5% were displayed three times across systems, and 7.4% were displayed more than three times across systems. Three titles were displayed as many as eight times across systems.

8.5.5 Resolving Conflicting Relevance Assessments

To compare the Blue and Pinstripe Systems using precision and estimated recall, we had to resolve conflicting relevance assessments. We reviewed relevance assessments for titles displayed more than once to determine how to resolve conflicting relevance assessments. We discussed user relevance assessments with project consultant Paul Kantor and showed him various reports on titles displayed more than one time. We noticed that the more users displayed the same titles, the more likely they were to give positive relevance assessments to titles, i.e., "very useful" or "possibly useful" ratings. After much discussion, we decided to use the last relevant assessment that searchers gave titles displayed more than once and to disregard other relevance assessments for the same title.

Two examples help to describe how we resolved conflicting relevance assessments. Let's say a user displayed a title twice in the Pinstripe System and gave it a "not useful" the first time she displayed it and a "possibly useful" rating the second time she displayed it. (The user did not display this title in her search of the Blue System.)

We kept the second rating, i.e., "possibly useful," and disregarded the first display and rating of this title in the Pinstripe System. Let's say a user displayed a title two times in the Pinstripe System and gave it a "not useful" rating followed by a "very useful" rating. The user then switched systems, displayed the very same title in a Blue-system search, and gave it a "possibly useful" rating. We kept the last rating, i.e., "possibly useful," converted the first rating of this title in the Pinstripe System to "possibly useful," and disregarded the second and third ratings of this title in the Pinstripe System.

In Table 8.9, the number of titles displayed in the Blue and Pinstripe Systems are given before and after the ASTUTE project team resolved conflicting relevance assessments.

Table 8.9. Titles Displayed Before and After Resolution of Conflicting Relevance Assessments

	UM-D				Earlham			
	No. in Blue System		No. in Pinstripe System		No. in Blue System		No. in Pinstripe System	
Search administration type	Titles displayed (before)	Titles displayed (after)	Titles displayed (before)	Titles displayed (after)	Titles displayed (before)	Titles displayed (after)	Titles displayed (before)	Titles displayed (after)
Search in one system only	294	286	319	304	85	74	51	50
Search in two systems, no or suspect questionnaire	519	430	711	653	90	85	101	80
Search in two systems, questionnaire	1,087	966	1,021	954	291	225	156	152
Total	1,900	1,682	2,051	1,911	466	384	308	282

The numbers of titles displayed before and after conflicting relevance assessments differed by as many as 218 titles (in UM-D's Blue System) and as few as 26 titles (in Earlham's Pinstripe System). Overall, conflicting relevance assessments affected less than ten percent of the titles that users displayed in the experimental systems.

Chapter 9 gives precision and estimated recall of experimental online catalog searches. To calculate these measures, we used the numbers given in the "Titles displayed (after)" columns of Table 8.9. Thus, precision and estimated recall scores discussed in this report were based on users' last relevance assessment of titles displayed in the last system they searched.

8.5.6 Rating Displayed Titles

Resolving conflicting relevance assessments required making the following changes to the relevance assessments users gave to displayed titles:

- Changing relevance assessments of titles in the first system searched to match the last relevance assessment users gave to titles in the second system searched.
- Disregarding multiple displays of the same title in the first system searched.
- Disregarding multiple displays of the same title in the second system searched.

When we made the above changes, there was one relevance assessment per unique title displayed in each of the two experimental systems. The one relevance assessment was the same across the two systems. If a user displayed the same title in both Blue and Pinstripe searches, there were two relevance assessments for the title — one assessment in the Blue System and one assessment in the Pinstripe System. Furthermore, the two assessments were the same and were based on the last relevance assessment users gave to the title in the second system searched.

Table 8.10 compares the number and percentage of titles users displayed in the Blue and Pinstripe Systems at UM-D and Earlham according to the three relevance assessments: (1) "not useful," (2) "possibly useful," and (3) "very useful." Percentages of titles in these three relevance assessment categories are given before and after conflicting assessments were resolved.

Table 8.10. Comparing Relevance Assessments Before and After Resolution of Conflicts

Institution/ system	No. of titles dis-played	No. Not useful	% Not useful	No. Possibly useful	% Possibly useful	No. Very useful	% Very useful
UM-D/ Blue (before)	1,900	434	22.8	234	12.3	1,232	64.9
UM-D/ Blue (after)	1,682	413	24.6	193	11.4	1,076	64.0
UM-D/ Pinstripe (before)	2,051	595	29.0	249	12.1	1,207	58.9
UM-D/ Pinstripe (after)	1,911	579	30.3	223	11.7	1,109	58.0
Earlham/ Blue (before)	466	204	43.8	92	19.7	170	36.5
Earlham/ Blue (after)	384	193	50.3	68	17.7	123	32.0
Earlham/ Pinstripe (before)	308	82	26.6	64	20.8	162	52.6
Earlham/ Pinstripe (after)	282	88	31.2	59	20.9	135	47.9

At UM-D, percentages of titles in the three relevance-assessment categories varied by less than 2% after conflicts were resolved. At Earlham, such percentages varied by as little as 0.1% and as much as 6.5%.

We compared sets of titles in each of the three relevance-assessment rating categories using t-tests to determine if the resolution of conflicting relevance assessments made a significant difference in the size of sets. The results of all t-tests were not significant. The t-test on titles rated "not useful" at Earlham came closest to producing significant results (p=0.0424). The closest any other test came to producing significant results was the t-test on titles rated "very useful" at Earlham (p=0.1100). Generally speaking, the resolution of conflicting relevance assessments had no significant effect on sets of rated titles.

8.6 Chapter Summary

Chapter 8 examined the characteristics of system users and searches. Characteristics of system users were gleaned from pre-search questionnaire responses (see section 8.2). Search characteristics came from an analysis of "usable" search administrations (see sections 8.3–8.5).

The majority of users at both UM-D and Earlham were searching ASTUTE for the first time (see figure 8.1). The majority were also daily users of computer systems. (see figure 8.2) Large percentages of UM-D respondents were computer science or engineering majors (29.1% and 46.1%, respectively). Thus, the majority of UM-D respondents were performing searches for topics in their major field of study. This was not true of respondents at Earlham College where ASTUTE's database focused on American history. Only one-quarter of Earlham respondents were history majors; consequently, the majority of Earlham respondents were not performing searches for topics in their major field of study (see Table 8.1).

Searchers did not spend much time using the ASTUTE experimental online catalog (see section 8.3). They spent less than a minute completing the pre-search questionnaire (see figures 8.3 and 8.4). Time spent searching was quite low for search administrations bearing suspect post-search questionnaires — a little more than three minutes at UM-D and two minutes at Earlham. Time spent searching was the longest for full search administrations — over four minutes at both UM-D and Earlham. Completing post-search questionnaires took less than a minute in search administrations with suspect questionnaire responses and took a minute and a half in full search administrations.

ASTUTE was designed to handle one query per search experiment administration (see section 8.4). Thus, the average number of queries per search (1.1 or 1.3) was much lower than the average number in operational online catalogs. A statistical test demonstrated that the length of queries had a significant effect on the ability of the experimental online catalog to make matches of its subject vocabulary (see Table 8.5). Long user queries (three or more words) were less likely to make matches than short queries (two or fewer words).

The first comparison of system performance was given in an analysis of system responses that designated matches of user queries (see Table 8.6). Matches gave users the opportunity to pursue the results of subject searching approaches. Generally, the Blue System was more likely to provide more opportunities for guiding users to useful information on their topics of interest than the Pinstripe System.

Providing users the opportunity to display titles did not guarantee that users would choose seemingly relevant subject headings from a list that would take them one step closer to retrieving titles connected with the subject heading. Furthermore, it did not guarantee that users would browse more than the first title displayed in free-text approaches. Users seldom displayed all the titles they retrieved, in fact, the statistics in Table 8.8 showed that they displayed less than 20% of the titles they retrieved. At UM-D, the Pinstripe System retrieved more than two times as many titles as the Blue System. High-posted searches in the Pinstripe System were likely to come from keyword-in-record searches; this search made retrievals based on matches in all subject-rich fields. Pinstripe-system users at UM-D displayed only about an eighth of the titles they retrieved.

At UM-D, Blue- and Pinstripe-system users displayed a total of 470 titles more than once. Of this total, 82.8% of titles were displayed twice across systems, 11.5% were displayed three times across systems, and 5.7% were displayed more than twice across systems. At Earlham, Blue- and Pinstripe-system users displayed a total of 183 titles more than once. Of this total at Earlham, 80.1% of titles were displayed twice across systems, 12.5% were displayed three times across systems, and 7.4% were displayed more than three times across systems. Section 8.5.5 tells how conflicting relevance assessments were resolved. The resolution of such conflicts did not have a significant effect on sets of rated titles. Resolution was necessary, however, to submit the data to statistical procedures to measure precision and estimated recall which was the focus of the data analyses in Chapter 9.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Markey, Karen. 1984. *Subject searching in library catalogs: before and after the introduction of online catalogs.* Dublin, Ohio: OCLC.

9 Precision and Estimated Recall in the Blue and Pinstripe Systems

9.1 Introduction

Chapter 9 features a discussion of relevance ratings in Blue- and Pinstripe-system searches and uses relevance ratings to measure and compare the search performance of the two experimental systems. This chapter also features end-user responses to the post-search questionnaire. Post-search questions probed user assessments of system performance, ease of system use, and importance of system features.

9.2 Precision in Experimental Catalog Searches

9.2.1 Overall Precision Scores

Precision scores were based on relevance assessments users gave to titles the last time they saw them in Blue- or Pinstripe-system searches. Precision scores for individual searches were calculated by dividing the number of titles retrieved and displayed and given "very useful" ratings in a search by the total number of titles retrieved and displayed in a search.

An example would help explain how precision is calculated. Let's say a user searched the Blue System first and displayed ten titles. She gave displayed titles five "not useful" ratings, three "possibly useful" ratings, and two "very useful" ratings. The user continued searching in the Pinstripe System and displayed six titles. She gave displayed titles two "not useful" ratings, two "possibly useful" ratings, and two "very useful" ratings. Two of the sixteen titles displayed were displayed more than once. One of the ten titles displayed in the Blue System was the same as one of the titles displayed in the Pinstripe System. The user rated this title "possibly useful" in the Blue System and "very useful" in the Pinstripe System. Since the user displayed this title last in the Pinstripe search, the relevance assessment for this title in the Blue System was changed to "very useful."

Of the six titles displayed in the Pinstripe System, one title was displayed twice. The first time the user displayed this title, she gave it a "not useful" rating; the second time the user displayed this title, she gave it a "possibly useful" rating. The relevance assessment for the first time this title was displayed was disregarded. Thus, this title retained the second relevance assessment for "possibly useful." Precision in the Blue System would be 30% (two "very useful" ratings and one "possibly useful" rating converted to "very useful" divided by ten displayed titles). Precision in the Pinstripe System would be 40% (two "very useful" divided by five displayed titles).

Figures 9.1 and 9.2 give average precision scores for the Blue and Pinstripe Systems according to search administration type at UM-D and Earlham.

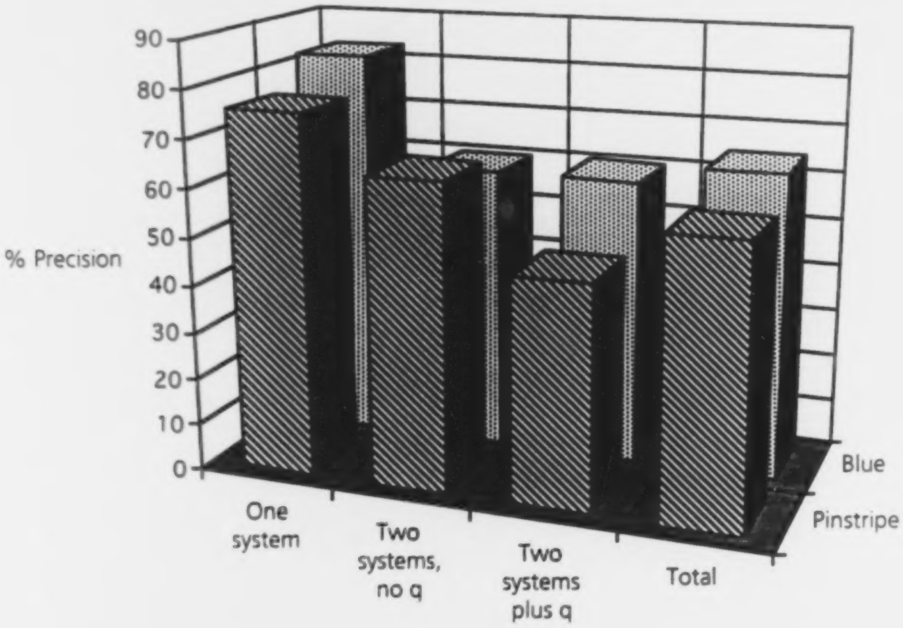


Figure 9.1. Precision by search administration type (UM-D)

At UM-D, precision scores averaged 64% in the Blue System and 58% in the Pinstripe System. Except for precision scores for search administrations in which users failed to complete questionnaires or gave suspect responses, precision scores in the Blue System exceeded precision scores in the Pinstripe System.

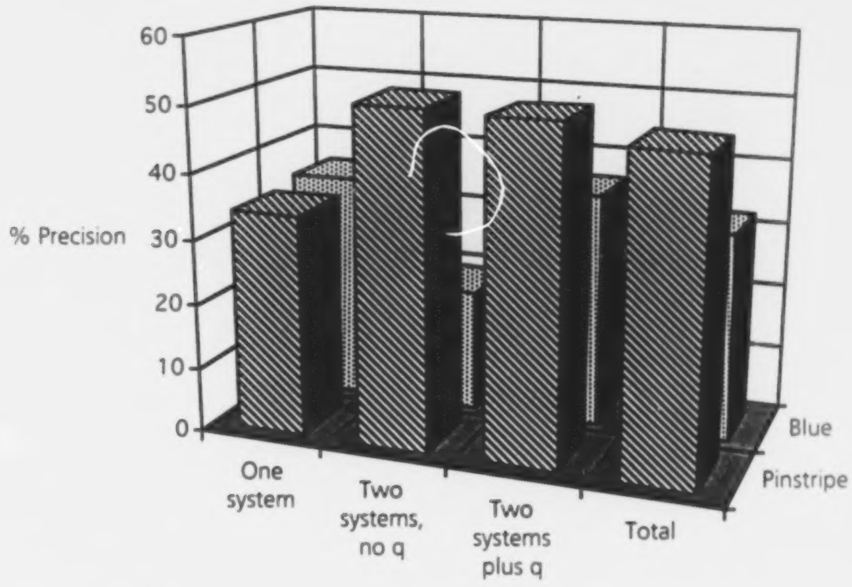


Figure 9.2. Precision by search administration type (Earlham)

The results were almost opposite at Earlham where precision scores averaged 32.0% in the Blue System and 47.9% in the Pinstripe System. The difference in precision scores was most evident for search administrations in which users failed to complete questionnaires or gave suspect responses; however, these precision scores were computed from "very useful" retrievals in 29 pairs of searches. The number of searches contributing "very useful" retrievals to "one-system" search administrations was 18 searches in the Blue System and 12 searches in the Pinstripe System. Thus, the numbers of cases for these two types of search administrations were quite small.

We compared precision scores for the two types of search administrations in which users conducted searches for the same topics in Blue and Pinstripe Systems to determine if the difference between scores was significant. The results of t-tests comparing precision scores are given in Tables 9.1 and 9.2. These tables limit t-test results to search administrations in which the same users conducted searches for the same topics in both Blue and Pinstripe Systems.

Table 9.1. T-test Results for Precision (UM-D)

Administration type	No. of titles displayed	Av. precision	Standard deviation
T-test results for two systems, no or suspect questionnaire			
Blue System	653	0.6432	0.4794
Pinstripe System	430	0.6000	0.4905
T = 1.4372	Degrees of freedom = 1,081		Probability = 0.1510
T-test results for two systems plus questionnaire			
No match	966	0.6014	0.4899
Successful match	954	0.4706	0.4994
T = 5.7929	Degrees of freedom = 1,918		Probability = 0.0000
Total (T-test results for two-system administrations)			
No match	1,396	0.6010	0.4899
Successful match	1,607	0.5408	0.4985
T = 3.3298	Degrees of freedom = 3,001		Probability = 0.0009

At UM-D, significant results were obtained for full search administrations and for the total of the two search administrations in which the same users conducted searches for the same topic in both Blue and Pinstripe Systems. Generally, precision scores in UM-D's Blue System were significantly greater than precision scores in UM-D's Pinstripe System.

At Earlham, significant results were obtained for all three analyses. Generally, precision scores in Earlham's Pinstripe System were significantly greater than precision scores in Earlham's Blue System.

Table 9.2. T-test Results for Precision (Earlham)

Administration type	No. of titles displayed	Av. precision	Standard deviation
T-test results for two systems, no or suspect questionnaire			
Blue System	85	0.1882	0.3932
Pinstripe System	80	0.5125	0.3932
T = -4.6282	Degrees of freedom = 163		Probability = 0.0000
T-test results for two systems plus questionnaire			
Blue System	225	0.3600	0.4811
Pinstripe System	152	0.5066	0.5016
T = -2.8524	Degrees of freedom = 375		Probability = 0.0046
Total (T-test results for two-system administrations)			
Blue System	310	0.3129	0.4644
Pinstripe System	232	0.5086	0.5010
T = -4.6929	Degrees of freedom = 540		Probability = 0.0000

9.2.2 Precision of Controlled Vocabulary Searches

The Blue and Pinstripe Systems featured controlled vocabulary and free-text approaches to subject searching. The Blue System's search trees favored controlled vocabulary approaches over free-text approaches in response to user queries. The Pinstripe System selected randomly amongst two controlled vocabulary and one free-text approaches.

Controlled vocabulary approaches to subject searching in the Blue System were: (1) exact, (2) alphabetical (for subjects and for personal names), (3) keyword-in-main-heading, and (4) keyword-in-subdivided-heading searches. The Pinstripe System featured the second and fourth controlled vocabulary approaches listed above. Free-text approaches to subject searching in the Blue System were: (5) title-keyword, (6) keyword in subject heading fields, and (7) keyword-in-record (for subjects and for personal names) searches. The Pinstripe System featured the seventh free-text approach listed above.

We categorized relevance ratings as to the particular approach (controlled vocabulary or free-text) responsible for their retrieval and display. We then calculated precision scores. Figures 9.3–9.5 compare precision scores achieved through controlled vocabulary searches to precision scores achieved through free-text approaches.

Figure 9.3 shows precision scores for UM-D's Blue System.

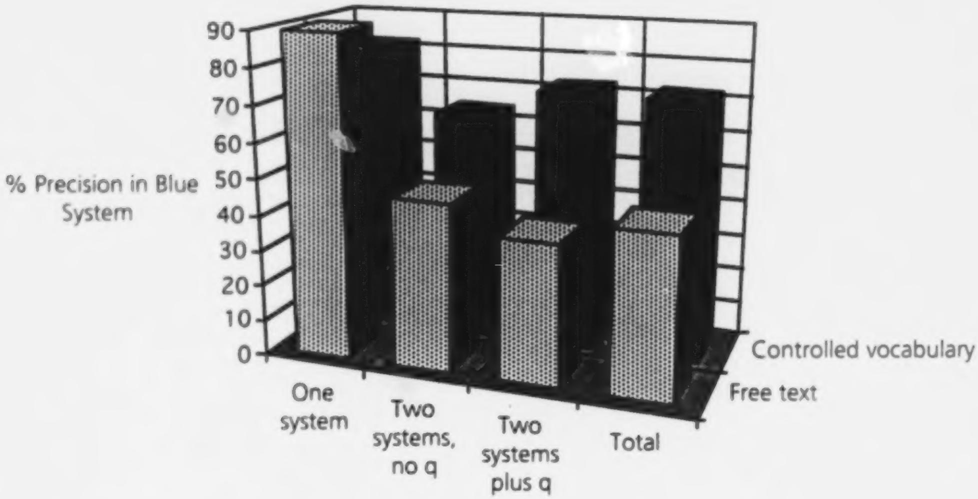


Figure 9.3. Comparing precision in Blue-system controlled vocabulary and free-text searches (UM-D)

In UM-D's Blue System, precision scores for controlled vocabulary approaches averaged 71.9% and precision scores for free-text approaches averaged 45.0%. Except for precision scores for one-system search administrations, precision in controlled vocabulary searches exceeded precision in free-text approaches.

We compared precision scores to determine if the difference between the scores for controlled vocabulary and free-text approaches was significant. The results of t-tests comparing precision scores in UM-D's Blue System are given in Table 9.3.

Table 9.3. T-test Results for Precision in the Blue System (UM-D)

Administration type	No. of titles displayed	Av. precision	Standard deviation
T-test results for Blue System, no or suspect questionnaire			
Controlled vocabulary	328	0.6402	0.4807
Free text	102	0.4706	0.5016
T = 3.0812	Degrees of freedom = 428		Probability = 0.0002
T-test results for Blue System plus questionnaire			
Controlled vocabulary	612	0.7222	0.4483
Free text	354	0.3927	0.4890
T = 10.6458	Degrees of freedom = 964		Probability = 0.0000
Total (T-test results for Blue System administrations)			
Controlled vocabulary	940	0.6936	0.4612
Free text	456	0.4101	0.4924
T = 10.5341	Degrees of freedom = 1,394		Probability = 0.0000

Significant results were obtained for all three analyses. Generally, precision scores for controlled vocabulary approaches in UM-D's Blue System were significantly greater than precision scores for free-text approaches in UM-D's Blue System.

Figure 9.4 shows precision scores for UM-D's Pinstripe System.

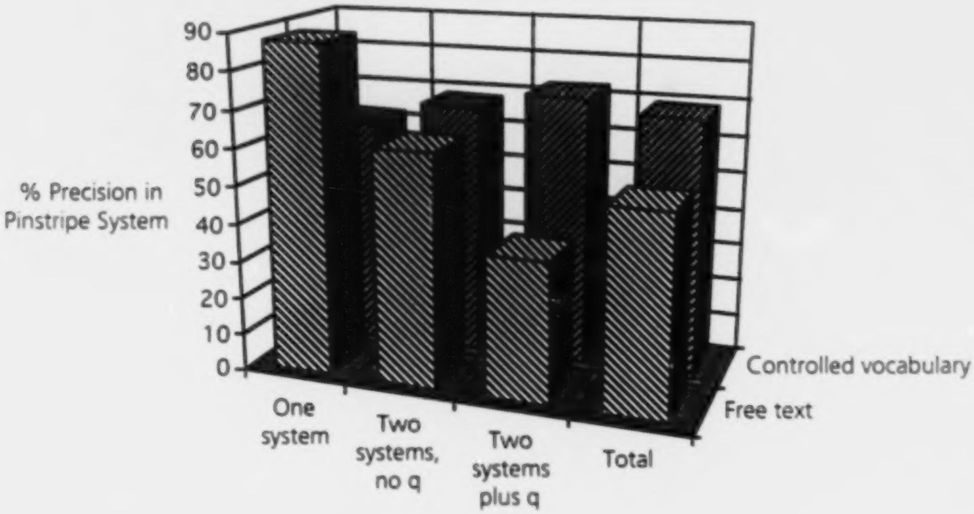


Figure 9.4. Comparing precision in Pinstripe-system controlled vocabulary and free-text searches (UM-D)

In UM-D's Pinstripe System, precision scores for controlled vocabulary approaches averaged 68.8% and precision scores for free-text approaches averaged 52.8%. Except for precision scores for search administrations in one system, precision in controlled vocabulary searches exceeded precision in free-text approaches.

We compared precision scores to determine if the difference between the scores for controlled vocabulary and free-text approaches was significant. The results of t-tests comparing precision scores in UM-D's Pinstripe System are given in Table 9.4.

Table 9.4. T-test Results for Precision in the Pinstripe System (UM-D)

Administration type	No. of titles displayed	Av. precision	Standard deviation
T-test results for Blue System, no or suspect questionnaire			
Controlled vocabulary	257	0.6770	0.4685
Free text	396	0.6212	0.4857
T = 1.4551	Degrees of freedom = 651		Probability = 0.1461
T-test results for Blue System plus questionnaire			
Controlled vocabulary	265	0.7245	0.4476
Free text	689	0.3730	0.4840
T = 10.2564	Degrees of freedom = 952		Probability = 0.0000
Total (T-test results for Blue System administrations)			
Controlled vocabulary	522	0.7011	0.4582
Free text	1,085	0.4636	0.4989
T = 9.1752	Degrees of freedom = 1,605		Probability = 0.0000

Significant results were obtained for two of the three analyses. Results were not significant for search administrations in which users failed to complete post-search questionnaires or gave suspect responses; however, precision scores for controlled vocabulary approaches were higher than such scores for free-text approaches. Generally, precision scores for controlled vocabulary approaches in UM-D's Pinstripe System were significantly greater than precision scores for free-text approaches in UM-D's Pinstripe System.

Figure 9.5 shows precision scores for Earlham's Blue System.

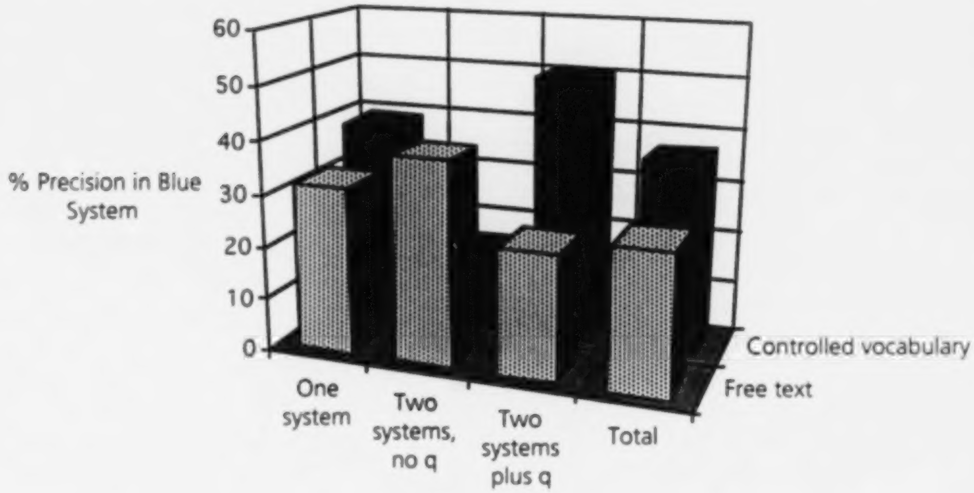


Figure 9.5. Comparing precision in Blue-system controlled vocabulary and free-text searches (Earlham)

In Earlham's Blue System, precision scores for controlled vocabulary approaches averaged 36.5% and precision scores for free-text approaches averaged 26.6%. Except for precision scores for search administrations in which users failed to complete post-search questionnaires or gave suspect responses, precision in controlled vocabulary searches exceeded precision in free-text approaches.

We compared precision scores to determine if the difference between the scores for controlled vocabulary and free-text approaches was significant. The results of t-tests comparing precision scores in Earlham's Blue System are given in Table 9.5. Significant results were obtained for only one analysis — for full search administrations.

Table 9.5. T-test Results for Precision in the Blue System (Earlham)

Administration type	No. of titles displayed	Av. precision	Standard deviation
T-test results for Blue System, no or suspect questionnaire			
Controlled vocabulary	72	0.1528	0.3623
Free text	13	0.3846	0.5064
T = -1.9907	Degrees of freedom = 83		Probability = 0.0498
T-test results for Blue System plus questionnaire			
Controlled vocabulary	103	0.5049	0.5024
Free text	122	0.2377	0.4274
T = 10.6458	Degrees of freedom = 223		Probability = 0.0000
Total (T-test results for Blue System administrations)			
Controlled vocabulary	175	0.3600	0.4814
Free text	135	0.2519	0.4357
T = 2.0433	Degrees of freedom = 308		Probability = 0.0419

Users conducted few searches in Earlham's Pinstripe System. Overall, precision for controlled vocabulary searches was 61.9% and precision for free-text searches was 45.4%. Table 9.6 gives the results of a single t-test for precision scores in search administrations in which users completed post-search questionnaires or gave suspect responses.

Table 9.6. T-test Results for Precision in the Pinstripe System (Earlham)

Administration type	No. of titles displayed	Av. precision	Standard deviation
Total (T-test results for Pinstripe System administrations)			
Controlled vocabulary	42	0.6190	0.4915
Free text	240	0.4542	0.4989
T = 1.9800	Degrees of freedom = 280		Probability = 0.0487

Results of the analysis were promising but not significant. For the most part, precision scores for controlled vocabulary approaches in Earlham's Pinstripe System were greater than precision scores for free-text approaches in Earlham's Pinstripe System.

With few exceptions, controlled vocabulary searches resulted in greater precision than free-text searches. This result supported the notion of search trees because the Blue System's search trees favored controlled vocabulary searches over free-text searches.

9.3 Estimated Recall Scores

Estimated recall scores were based on relevance assessments users gave to titles the last time they saw them in Blue- and Pinstripe-system searches. Recall scores for individual searches were calculated by dividing the number of titles retrieved and displayed and given "very useful" ratings in a search by the total number of titles retrieved and displayed and given "very useful" ratings" in the two searches the user conducted for his topic in the Blue and Pinstripe Systems.

Let's use the same example as we used in section 9.3.1 in a discussion of calculating precision to explain how recall was calculated. The user searched the Blue System first and displayed ten titles. He gave displayed titles five "not useful" ratings, three "possibly useful" ratings, and two "very useful" ratings. The user continued searching in the Pinstripe System and displayed six titles. She gave displayed titles two "not useful" ratings, two "possibly useful" ratings, and two "very useful" ratings. Two of the sixteen titles displayed were displayed more than once. One of the ten titles displayed in the Blue System was the same as one of the titles displayed in the Pinstripe System. The user rated this title "possibly useful" in the Blue System and "very useful" in the Pinstripe System. Since the user displayed this title last in the Pinstripe search, the relevance assessment for this title in the Blue System was changed to "very useful." Of the six titles displayed in the Pinstripe System, one title was displayed twice. The first time the user displayed this title, she gave it a "not useful" rating; the second time the user displayed this title, she gave it a "possibly useful" rating. The relevance assessment for the first time this title was displayed was disregarded. Thus, this title retained the second relevance assessment for "possibly useful."

After conflicting relevance assessments were resolved, the search in the Blue System resulted in five titles deemed "not useful," two titles deemed "possibly useful," and three titles deemed "very useful;" the search in the Pinstripe System resulted in one title deemed "not useful," two titles deemed "possibly useful," and two titles deemed "very useful." The number of titles given "very useful" ratings was four because one of the titles deemed "very useful" was displayed in both systems. Thus, recall in the Blue System was 75% (3 "very useful" titles displayed in the Blue System divided by the total of 4 titles deemed "very useful" in searches of the two experimental systems) and recall in the Pinstripe System was 50% (2 "very useful" titles displayed in the Pinstripe System divided by the total of 4 titles deemed "very useful" in searches of the two experimental systems).

Recall scores reported in this study were estimates of this measure. To figure recall scores that were not estimates, the researchers would have to perform a comprehensive review of every title in the two experimental online catalogs for every search topic and enlist subject experts for relevance assessments (Lancaster and Warner 1993, 181). Such a review was not within the purview of this study. The method of calculating recall employed in this study has been suggested as a way of estimating recall (Lancaster and Warner 1993, 181–3).

Figure 9.6 gives recall scores for experimental system users at UM-D who conducted searches in both Blue and Pinstripe Systems. This figure does not include scores of users who conducted a search of one of the two experimental systems because there was no second search for the same topic from which to extract relevance ratings for the calculation of estimated recall scores.

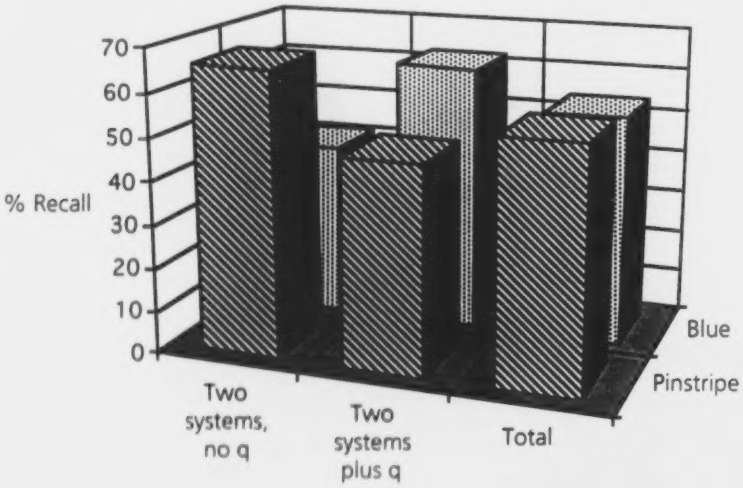


Figure 9.6. Estimated recall (UM-D)

Overall, recall scores in UM-D's experimental online catalog were about the same; recall in the Blue System was 53.4% and recall in the Pinstripe System was 55.4%. Recall scores in the two search administrations types flip flopped.

Figure 9.7 gives recall scores for experimental system users at Earlham who conducted searches in both Blue and Pinstripe Systems. This figure does not include scores of users who conducted a search of one of the two experimental systems because there was no second search for the same topic from which to extract relevance ratings for the calculation of estimated recall scores.

Overall, recall in Earlham's Pinstripe System (61.8%) was somewhat higher than recall in the Blue System (50.8%). Recall scores in the two search administration types flip flopped. Recall scores (74.5%) in the Pinstripe System for search administrations in which users did not complete post-search questionnaires or gave suspect responses were considerably higher than such scores in the Blue System (29.1%); however, recall scores for this search administration type were based on only eighteen search administrations.

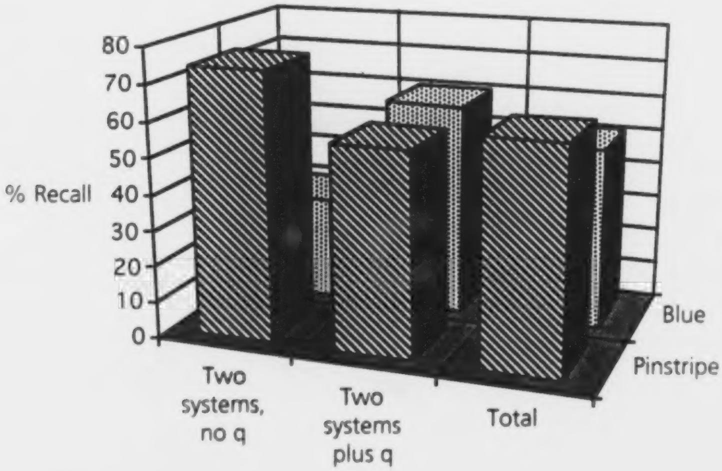


Figure 9.7. Estimated recall (Earlham)

Estimated recall scores varied greatly depending on the type of search administration. In the most complete search administrations (when users completed both searches in the Blue and Pinstripe Systems and the post-search questionnaire), recall scores were higher in the Blue System than in the Pinstripe System.

We did not perform t-tests on recall scores because t-tests, by definition, are used to determine whether the means of two *independent* samples differ significantly (Roscoe 1975, 224). In online retrieval tests, data were gathered and manipulated in such a fashion that the two samples were not independent.

The analysis of rated titles enlisted a relative recall calculation that required adding the total number of relevant documents found in one system to the total number of relevant documents found in a second system. For this experiment, the relative recall was calculated by dividing the number of titles rated "very useful" in the Blue or Pinstripe System by the total number of titles rated "very useful" in both searches for the same topic in the Blue and Pinstripe Systems.

System users performed searches and made relevance assessments on the titles they displayed. Ideally, users performed relevance assessments of titles retrieved and displayed in both Blue and Pinstripe Systems. Users occasionally rated titles retrieved in both Blue and Pinstripe Systems. Because titles could be displayed and rated in both systems, the samples were not independent. Thus, a t-test of two independent samples was not a valid statistical test for these data.

9.4 Post-search Questionnaire Responses

9.4.1 System Performance

Of the eleven questions in the post-search questionnaire, three questions specifically asked end users about the performance of the two experimental systems in finding useful titles for their topics of interest. These questions asked users to compare the two experimental systems in terms of the number of useful titles they retrieved, their satisfaction with the two systems, and their system preference for future searches.

Users who completed post-search questionnaires were a subset of the users who searched the experimental online catalogs and whose precision and estimated recall scores were given in several tables and figures of section 9. Users who completed post-search questionnaires searched both Blue and Pinstripe Systems to find titles on their topics of interest. In table and figure captions, precision and estimated recall scores that these users' searches achieved were designated by the caption "Search in two systems, questionnaire" or "Two systems plus q." Totals of 171 users at UM-D and 57 users at Earlham answered post-search questions.

The first post-search question asked users to compare the two systems based on the number of useful titles they retrieved. The first question and four responses categories were:

"Comparing the Blue Test System with the Pinstripe Test System, did you find:

- A greater number of useful titles in the Blue Test System
- About the same number of useful titles in both systems
- A smaller number of useful titles in the Blue System
- No useful titles in either system

Figure 9.8 summarizes end-user responses in search administrations at UM-D and Earlham.

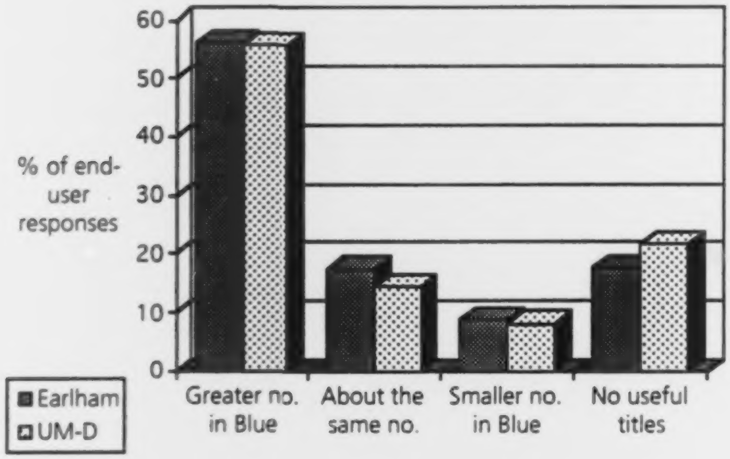


Figure 9.8. Comparing the number of useful titles retrieved

Almost three-quarters of users felt that the Blue System retrieved a greater number of useful titles than or about the same number of useful titles as the Pinstripe System.

The second post-search question asked users about their satisfaction with their searches. This question and three responses categories were:

"In relation to what you were looking for, would you say your search was more satisfactory in:

- Blue Test System
- Pinstripe Test System
- No difference

Figure 9.9 summarizes end-user responses to this question at UM-D and Earlham.

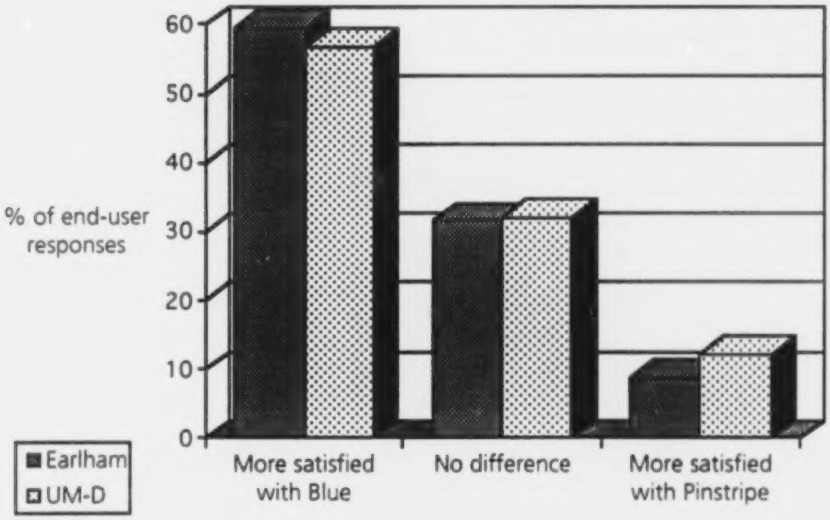


Figure 9.9. Satisfaction with the search

The vast majority of end users felt their search was more satisfactory in the Blue System than in the Pinstripe System. At UM-D, 56.7% of end users chose the Blue System. At Earlham, 59.6% of end users chose the Blue System. About a third of users at both institutions felt no difference in satisfaction with their search in the two experimental systems. The percentages of users who were more satisfied with their search of the Pinstripe System were quite small (12.3% of UM-D users and 8.8% of Earlham users). The low percentage of Earlham users who were more satisfied with their search in the Pinstripe System was rather surprising in view of the higher overall precision scores for searches of the Pinstripe System than for searches of the Blue System (see figure 9.2). However, other precision scores for Pinstripe-system searches were lower than such scores for Blue-system searches (see Table 9.6).

The third post-search question asked users which system they preferred. This question and three responses categories were:

"The next time you need to conduct a computer search, will you want to use:

- Blue Test System
- Pinstripe Test System
- No difference

BEST COPY AVAILABLE

Figure 9.10 summarizes end-user responses to this question on system preference.

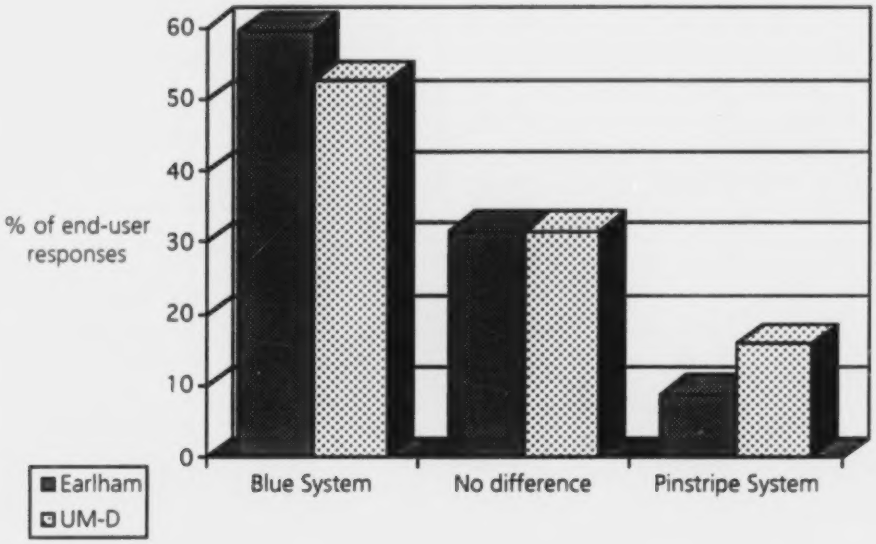


Figure 9.10. System preference

The majority of end users preferred the Blue System over the Pinstripe System. At UM-D, 52.6% of end users chose the Blue System. At Earlham, the percentages were unchanged from the previous question. Clearly, end users at Earlham preferred the Blue System. About a third of users at both institutions did not have a preference for one system over another.

9.4.2 Ease of System Use

Measures such as precision and estimated recall that information retrieval researchers have used over the years to evaluate systems were important to the end users who participated in the retrieval tests in this study because their responses to post-search questions regarding retrieving useful titles, search satisfaction, and system preference favored the Blue System which gave them high precision and recall. The post-search questionnaire probed other factors — ease of system use, system efficiency in retrieving useful titles, system helpfulness in giving users new subject searching ideas — in the comparison of Blue and Pinstripe Systems.

The next five questions on the post-search questionnaire dealt with ease of system use. The first question asked users to compare the ease of giving instructions to the two systems. The question was "Which system do you find easier to give instructions to, for example, to type in subjects, continue searches, make selections from lists?"

Figure 9.11 summarizes end-user responses to this question.

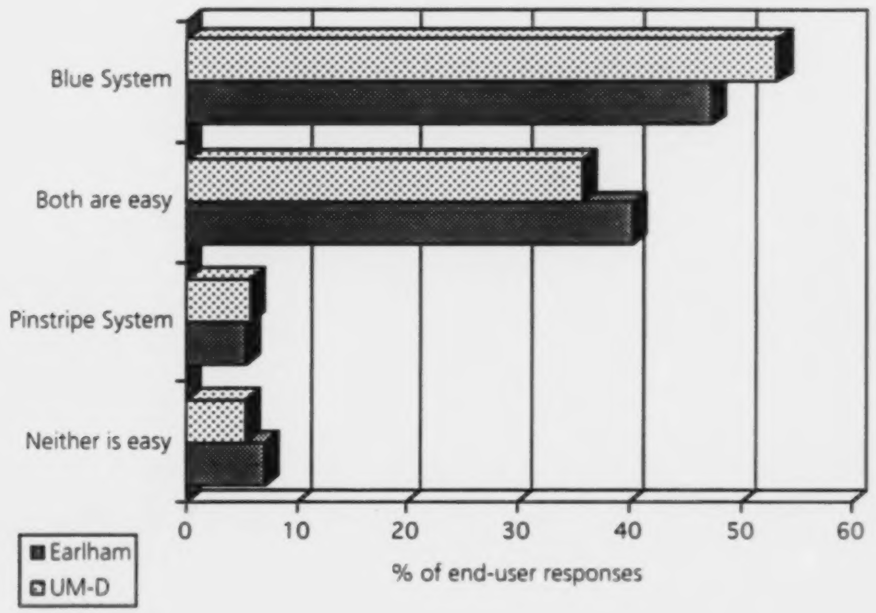


Figure 9.11. Ease of giving instructions to the experimental systems

Generally, users at both institutions chose the Blue System over the Pinstripe System as the easier of the two systems to which to give instructions. Large percentages of users (35.7% and 40.3% of UM-D and Earlham users, respectively) felt both systems were easy to give instructions to.

Responses to the next four questions on ease of system use fell into the same pattern, that is, large percentages of users chose the Blue System over the Pinstripe System and sizable percentages of users felt positively about both systems about the particular ease-of-use aspect under scrutiny.

Figure 9.12 summarizes user responses to a question about getting instructions from the system, viz. "Which system gives you more useful instructions in response to the subjects you type in?"

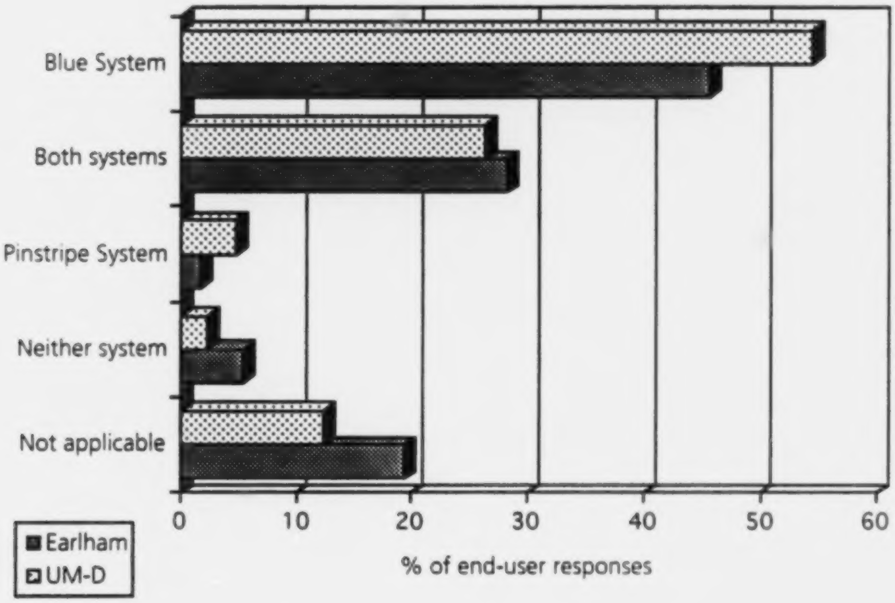


Figure 9.12. Ease of getting instructions from the experimental systems

Besides fitting into the response pattern described above, between 12% and 20% of users selected the "not applicable" category. Many of these users had conducted searches for subjects that failed to produce retrievals. Following a system message informing users that the system found nothing in the system, the system continued with the post-search questionnaire. Thus, the experimental systems gave users little or no opportunity to assess whether system instructions were useful in response to user queries because they retrieved no titles for user queries.

Figure 9.13 summarizes user responses to a question about the systems' clarity, viz. "Which system tells you what to do more clearly?"

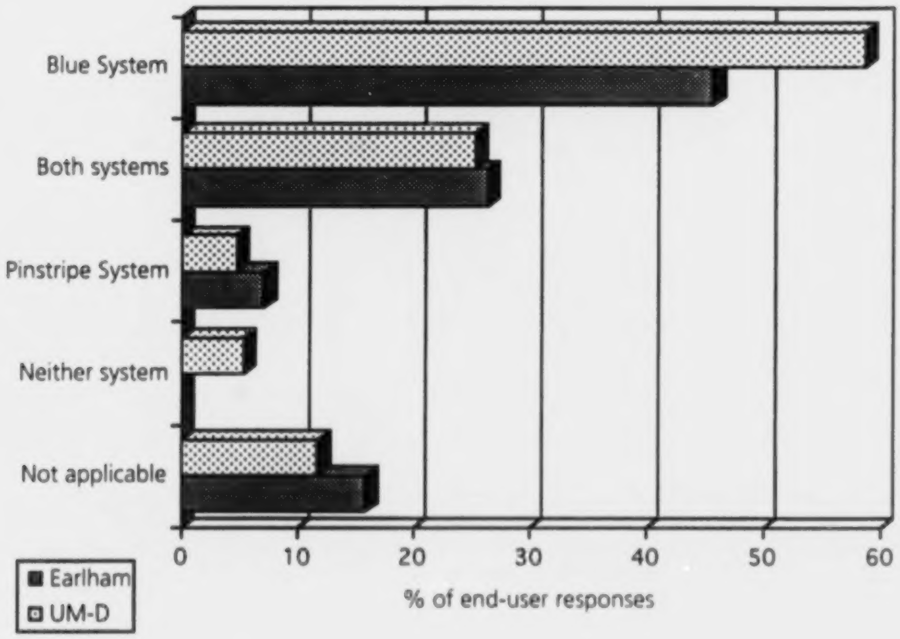


Figure 9.13. Clarity of the experimental systems

User responses followed the same pattern as the previous question. That is, many users chose the Blue System over the Pinstripe System in terms of clarity. About a quarter of users felt both systems were clear. Sizable percentages of users (11.7% at UM-D and 15.8% at Earlham did not feel the question was applicable to their experience using the experimental systems. The reason could have been attributed to searches in which no titles were retrieved because the systems did not give users the opportunity to assess their respective clarity.

Figure 9.14 summarizes user responses to a question about the systems' efficiency, viz. "Which system is more efficient in guiding you to useful information?"

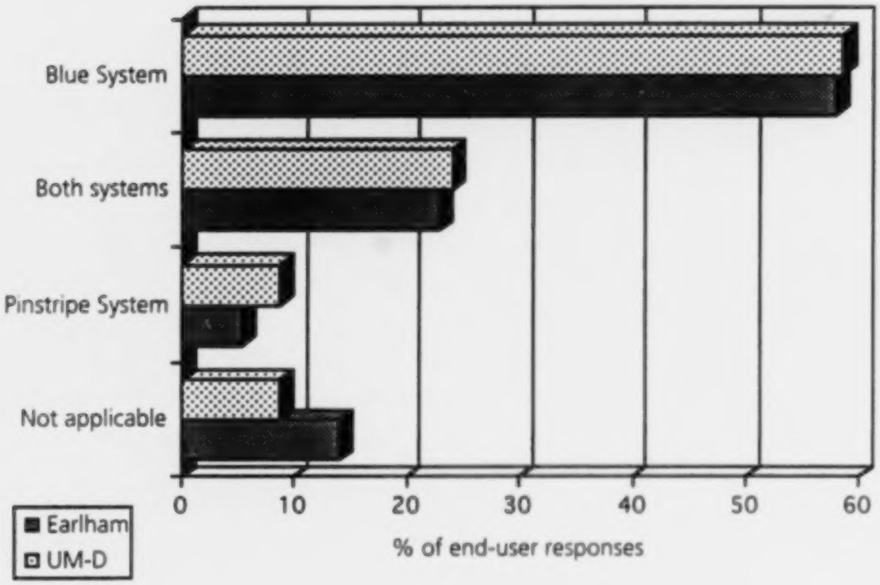


Figure 9.14. Efficiency of the experimental systems

The familiar pattern of favoring the Blue over the Pinstripe System was evident in user responses to the question on efficiency. About a quarter of respondents felt both systems were efficient in guiding them to useful information

Figure 9.15 summarizes user responses to a question about the systems' helpfulness in giving users new ideas for subject searching, viz. "Which system is more helpful for giving you new ideas for searching for subjects?"

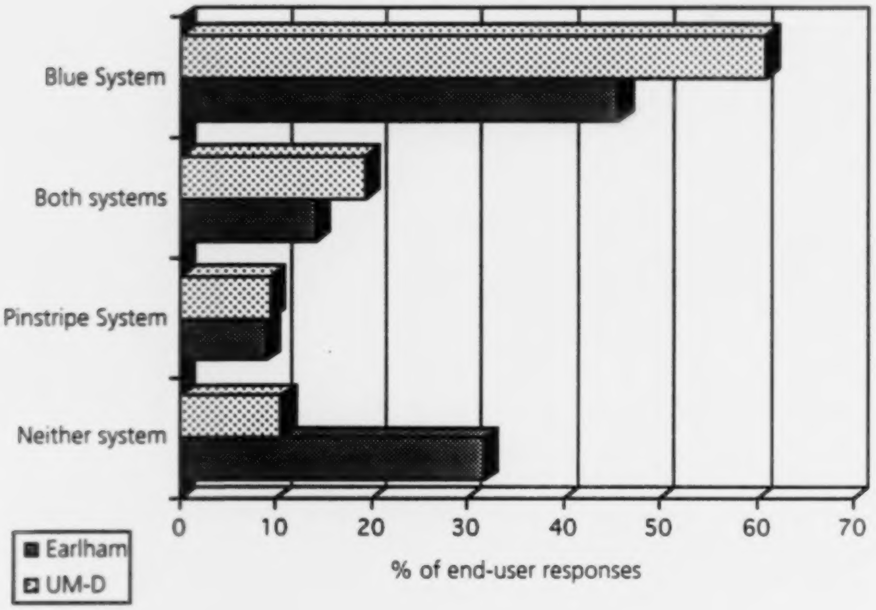


Figure 9.15. Helpfulness of giving new ideas for subject searching

The response pattern was a little different for this question. Respondents at UM-D (60.8%) clearly favored the Blue System over the Pinstripe System. Only about ten percent of UM-D users felt neither system was helpful in giving new ideas for subject searching. Respondents at Earlham (45.6%) were somewhat less exuberant about the helpfulness of the Blue System. Almost a third of Earlham users felt neither system was helpful in this regard. Recall the lower estimated precision score that the Blue System achieved at Earlham (figure 9.2). Although Earlham users clearly favored the Blue System overall, their response to this question indicated that both systems need improvement in capabilities that help users get new ideas for subject searching.

9.4.3 Importance of System Features

Following the five questions on ease of system use, the questionnaire enumerated six system features and asked users to tell us how important each feature was to them. The question was "On a scale of 1 to 5 in which 1 means 'not at all important' and 5 means 'very important,' rate the following features as to their importance in helping you search library computer systems like ASTUTE." The system prompted users to enter a number between 1 and 5. It listed six features for users to rate. The features were:

- System's ability to continue searching on its own to find useful information
- System's efficiency in guiding you to useful information
- System's clarity in telling you what to do
- New ideas that the system gives you for searching for subjects
- Usefulness of the system's instructions
- Ease of giving instructions to the system

All but the first-listed feature above were similar to the ease-of-use aspects that had been addressed in the previous five questions.

We included this question to determine what system features were especially important to users. Figure 9.16 shows the percentage of users who rated features "important" or "very important."

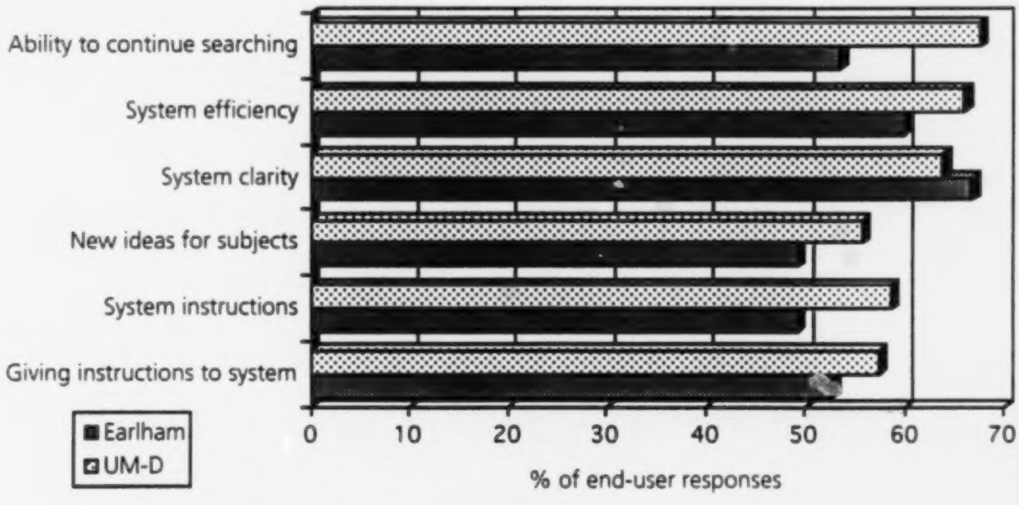


Figure 9.16. Importance of system features

End users at both institutions were positive about all listed features. End users at UM-D were especially keen on system efficiency and a system's ability to continue searching on its own to find useful information. End users at Earlham were more positive about system clarity and efficiency.

9.4.4 Concluding Questions

The post-search questionnaire concluded with two questions. The first questioned users about their familiarity with the subject they searched. The question was "How would you rate your familiarity with the subject you searched today?" Figure 9.17 summarizes end-user responses at UM-D and Earlham.

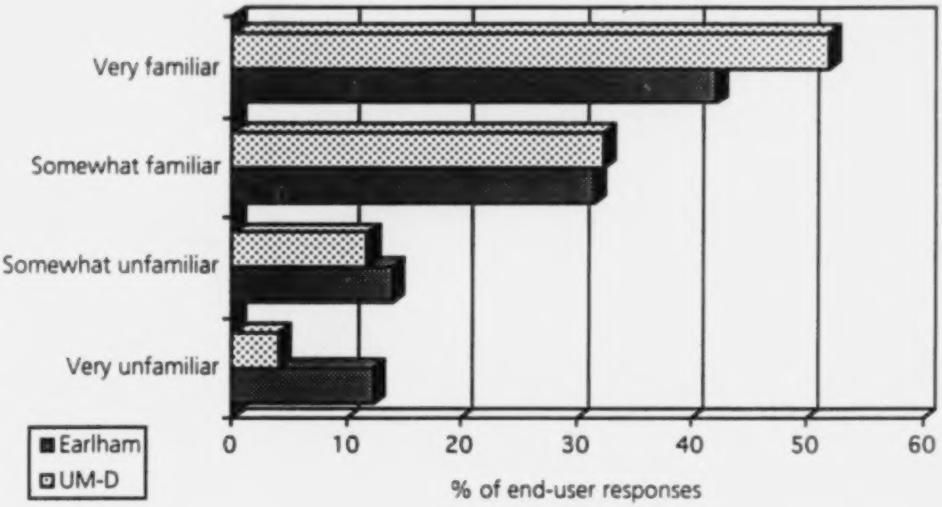


Figure 9.17. User familiarity with subject searched

Large percentages of users were very familiar with the subject they searched in the experimental online catalog. Although Earlham users reported a lesser degree of familiarity than UM-D users, almost three-quarters of Earlham users said they were "very familiar" or "somewhat familiar" with their topics of interest.

Only a quarter of Earlham users reported history as their major field of study (see Table 8.1). The experimental online catalog's database contained library material on American history. Thus, the lesser degree of familiarity reported by Earlham users was expected in view of the major fields of study they reported.

The final question asked users how interested they were in seeing the capabilities of the experimental online catalogs extended to topics besides Engineering and Computer

Science (at UM-D) and American history (at Earlham). Figure 9.18 gives user responses to the question.

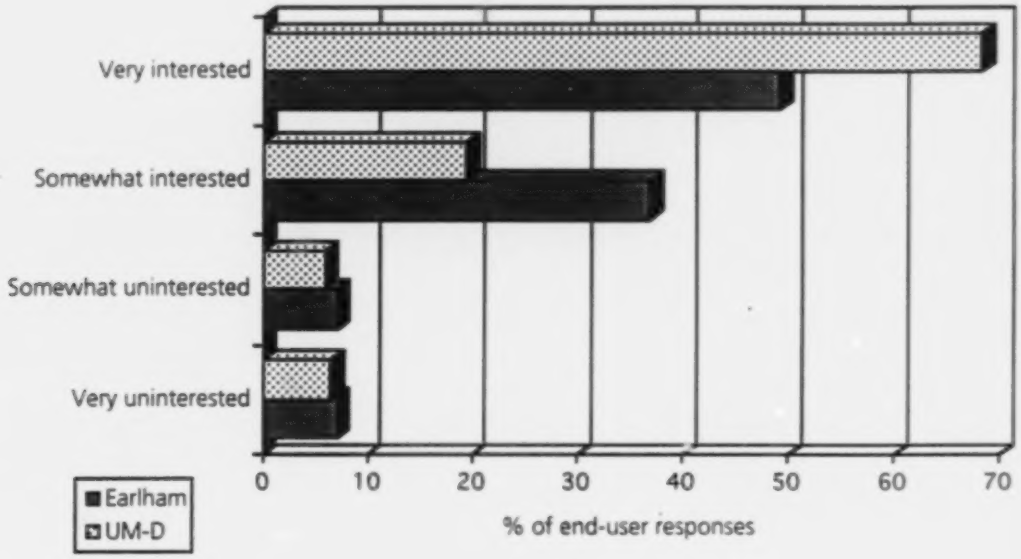


Figure 9.18. Extending ASTUTE features to other topics

We interpreted positive responses to this question as an indication that users were generally pleased with the capabilities of the experimental ASTUTE system. Almost seven of every eight users at UM-D were interested in seeing ASTUTE features extended to other topics. Earlham users were less enthusiastic than UM-D users but they were still quite positive. ASTUTE prompted users who selected the "very interested" response category to type in the topics to which they would ASTUTE features extended. Table 9.7 lists user-suggested topics.

Table 9.7. Topics to Extend to ASTUTE Features

Topics Suggested by UM-D Users	Topics Suggested by Earlham Users
Business	Art
Politics	Art history
Arts and humanities	Philosophy
Physics	Sociology
Biology	Social sciences
Psychology	World War 2
More than just engineering	African history
Natural sciences	Civil War
Geometry	Politics
History	African American poetry
Foreign languages	Religion
Law	Child abuse
Sociology	Supreme Court decisions
Flowcharting	African American women writers in Harlem Renaissance
Digital signal processing	Women in Reconstruction
Solar powered automobiles	
Object oriented programming	

Many suggested topics named broad fields of study, e.g., geometry, physics, biology. Some topics described specific subjects that had limited coverage in the UM-D and Earlham databases, e.g., "object oriented programming," "digital signal processing," "African American women writers in Harlem Renaissance," and "Women in Reconstruction."

9.5 Chapter Summary

This chapter used the traditional measures of precision and estimated recall to compare the performance of the Blue and Pinstripe Systems. Statistical results were supported by end-user responses to questions in post-search questionnaires.

At UM-D, precision scores averaged 64% in the Blue System and 58% in the Pinstripe System. In full search administrations at UM-D (i.e., administrations in which users conducted searches in both Blue and Pinstripe Systems and completed a post-search questionnaire), precision scores in the Blue System were significantly higher than precision scores in the Pinstripe System (see figure 9.1. and Table 9.1).

The result was exactly the opposite at Earlham. That is, precision scores in the Pinstripe System were significantly higher than precision scores in the Blue System (see figure 9.2. and Table 9.2).

We extended the analysis of precision scores to comparisons for controlled vocabulary and free-text searches because the Blue System's search trees favored controlled vocabulary approaches over free-text approaches in response to user queries. We expected the precision of titles retrieved through controlled vocabulary approaches to exceed the precision of titles retrieved through free-text approaches. Such results would support the design of search trees in the experimental online catalog.

In UM-D's Blue and Pinstripe Systems, precision scores for controlled vocabulary approaches exceeded such scores for free-text approaches (see figures 9.3 and 9.4). Furthermore, the differences between scores were significant (see Tables 9.3 and 9.4). Especially large differences between precision scores were registered for full search administrations. In Earlham's Blue System, precision scores for controlled vocabulary approaches exceeded such scores for free-text approaches in full search administrations (see figure 9.5). There was an insufficient number of searches for statistical comparisons of precision in Earlham's Pinstripe System; however, overall precision scores for controlled vocabulary approaches were higher than such scores for free-text approaches at Earlham (see Table 9.6). These findings about precision and search approaches support the design of search trees that favor controlled vocabulary over free-text approaches for searching bibliographic databases.

Overall, recall scores in UM-D's experimental online catalog were about the same; recall in the Blue System was 53.4% and recall in the Pinstripe System was 55.4% (see figure 9.6). Recall scores in the two search administrations flip flopped. That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Type of search administration results were the same at Earlham — recall scores in the two search administrations flip flopped (see figure 9.7). That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Overall, recall in Earlham's Pinstripe System (61.8%) was somewhat higher than recall in the Blue System (50.8%). No statistical tests were performed on recall scores because the design of the comparison search experiment violated the t-test's principle of independence.

End-user responses to the post-search questionnaire were available for full search administrations. Generally, responses reflected statistical results of this type of search administration. That is, users rated the performance of the Blue System higher than the Pinstripe System in terms of retrieving greater numbers of useful titles (see figure

9.8), search satisfaction (see figure 9.9), and they preferred the Blue System over the Pinstripe System (see 9.10).

The post-search questionnaire probed other factors — ease of system use, system efficiency in retrieving useful titles, system helpfulness in giving users new subject searching ideas — in the comparison of Blue and Pinstripe Systems. The majority of end users rated the Blue System or both systems highly in terms of the ease of giving instructions to the systems (see figure 9.11), ease of getting instructions from the system (see figure 9.12), system clarity (see figure 9.13), and system efficiency in guiding them to useful information (see figure 9.14).

At Earlham, users were a little less enthusiastic about the Blue System's helpfulness in giving new ideas for subject searching (see figure 9.15). Since the design of the Blue System was intended to give users new subject searching ideas through exact-approach browsing and expanding opportunities, we were somewhat concerned about this response. The failure analysis of experimental catalog searches examines searches closely to determine the reasons for failure.

References

Lancaster, F. Wilfred, and Amy J. Warner. 1993. *Information retrieval today*. Arlington, Va.: Information Resources Press.

Roscoe, John T. 1975. *Fundamental research statistics for the behavioral sciences*. New York: Holt, Rinehart, and Winston.

10 Failure Analysis of User Searches

10.1 Introduction

Recall and precision are measures that information retrieval researchers have used for many years to evaluate searches (Lancaster and Warner 1993). While these measures disclose important information about the number of useful titles searches produce, they do not reveal why systems are successful or unsuccessful in retrieving useful titles.

This and subsequent chapters feature a failure analysis of end-user searches to determine why the experimental ASTUTE online catalog was successful or unsuccessful in retrieving useful titles. The project director, an expert searcher of information retrieval systems, repeated end-user searches that were successful and unsuccessful in retrieving useful titles to identify the reasons why they succeeded or failed. This section features the results of the failure analysis.

10.2 Successful and Unsuccessful Searches at UM-Dearborn

The first step in the failure analysis of searches performed at the University of Michigan-Dearborn and Earlham College was to characterize searches that users performed in the Blue and/or Pinstripe Systems as successful or unsuccessful:

- Successful search — users rated one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful search — users did not rate one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful pairs of searches — users did not rate titles "very useful" in searches performed in the Blue *and* Pinstripe Systems for the *same* topic.

Table 10.1 enumerates the number and percentage of successful and unsuccessful searches in the Blue and Pinstripe Systems at UM-D.

Table 10.1. Successful and Unsuccessful Searches at UM-D

Useful titles retrieved	Blue System		Pinstripe System	
	No.	%	No.	%
≥1 useful titles retrieved	156	44.2	150	42.1
No useful titles retrieved in one of two systems	69	19.5	78	21.9
No useful titles retrieved in searches of both systems	128	36.3	128	36.0
Total	353	100.0	356	100.0

Percentages of successful and unsuccessful searches were about the same for the Blue and Pinstripe Systems. One or more useful titles were retrieved in 44.2% of searches in the Blue System; one or more useful titles were retrieved in 42.1% of searches in the Pinstripe System. The Blue System had a lower percentage (19.5%) of searches in which no useful titles were retrieved than in the Pinstripe System (21.9%) but the percentages were not much different from one another. Generally, the Blue and Pinstripe Systems performed about the same in terms of retrieving useful titles in searches.

Subsections of section 10.2 focus on unsuccessful searches in the Blue and Pinstripe Systems at UM-D.

10.2.1 Failure of Both Systems

At UM-D, users performed a total of 297 pairs of searches. That is, they performed one search for their topic of interest in the Blue Systems followed by a search for the same topic in the Pinstripe System. Of the 297 pairs, 128 search pairs failed to yield useful titles in searches of both experimental systems — about 43% of the search pairs. The project director, an expert searcher of information retrieval systems, repeated these 128 searches in both Blue and Pinstripe Systems to determine why end users were unsuccessful in retrieving useful titles.

10.2.1.1 User Perseverance

User perseverance was a common reason for the failure of searches to retrieve useful titles. For example, a user entered the subject query "fiber optics" into the Pinstripe System. The keyword-in-record search responded by displaying the first of fifteen retrieved titles to the user. The user gave this first title a "not useful" rating and terminated the search. Examples of titles this user missed by failing to continue the display of retrieved titles were *Fiber optics engineering, Fundamentals of optical fiber communication, Fiber optics,* and *Optical fiber systems: technology, design, and* *applications.* This and many other searches like it should not be counted as unsuccessful searches. The catalog produced titles on the subject query that the user entered. The user chose to terminate his search rather than display additional titles.

Perseverance was also the reason why the user's search of the Blue System for the same topic failed to produce useful titles. In response to the query "fiber optics," the Blue System made an exact match of a subject heading and responded with exact-search options for displaying general works, browsing broader terms, displaying form subdivisions, and expanding the search using a different subject searching approach. Instead of pursuing one or more of the exact-search options, the user terminated the search.

Table 10.2 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to user perseverance.

Table 10.2. User Perseverance

User query	Explanation
ceramics	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, notes, form subdivisions, and for expanding the search are enabled. The user ends the search.
probability	The Blue System's title-keyword search displays the first of 24 retrieved titles. The user gives it a negative relevance assessment and ends the search. By failing to display additional titles, the user misses potentially useful titles such as *Probability and statistics for engineers and scientists, Probability and statistics for engineers,* and *Probability and statistics for modern engineering.*
operations research	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, subject and form subdivisions, and for expanding the search are enabled. The user ends the search. Had she selected the general works option, she would have retrieved titles such as *Operations research: An introduction, Introduction to operations research,* and *Applied operations research and management science.*
chemical engineering	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, subject and form subdivisions, and for expanding the search are enabled. The user selects the general works option. The system responds by displaying the first of 32 retrieved titles. The user rates this title "not useful" and ends the search.

If users had continued searching for the searches described in Table 10.2, these searches would have most likely succeeded in retrieving useful titles.

The researchers did not conduct personal interviews with end users following their search(es). Had we conducted such interviews, we would have had explanations from users as to their reasons for terminating their searches. In the absence of interview data, we can speculate that the results of alphabetical or exact searches could have been too complex for users to understand. That is, they could have been lost or confused about which of the several options to pursue for continuing the search. This does not explain what was in the mind of the user who conducted the search for "chemical engineering." The Blue System retrieved a manageable number of titles (32) and the user failed to display more than the first title.

In most controlled vocabulary searches that failed due to perseverance, users failed to make actions (e.g., select an exact-search option, select a subject heading from an alphabetical list) that would enable them to retrieve and display titles. In several free-text searches, users failed to display more than the first of several retrieved titles.

Users conducted two-thirds of the 39 searches that failed due to perseverance in the Blue System. In several searches, users made a few actions to find useful titles. Two examples of such searches in the Blue and Pinstripe Systems follow. One user entered the query "c programming language" into the Pinstripe System which responded with the alphabetical approach and placed the user in an alphabetical list bearing subject headings such as "C (Computer program language)," "C Basic (Computer program language)," and "C++ (Computer program language)." The user selected the former, gave two of 42 retrieved titles negative relevance assessments, and ended the search.

Another user entered the query "mathematics" into the Blue System which made an exact match of a subject heading and responded with the exact approach. The user browsed narrower terms, selected the narrower term "Logic, Symbolic and mathematical," displayed three of nine retrieved titles, and gave them negative relevance assessments. The user had several options for continuing the search. He could have browsed titles assigned to subdivided forms of the selected narrower term; he could have returned to the exact-approach main menu for the original matching subject heading "Mathematics" and browsed titles assigned to this or subdivided forms of this subject heading. Instead, the user terminated the search without browsing or exploring terminology or titles connected to mathematics.

The failure analysis of searches that failed due to user perseverance demonstrates that these searches were really successful searches in terms of leading users to titles on their topics of interest. Despite the several opportunities the systems gave to users for browsing terminology or displaying titles, users terminated their searches.

10.2.1.2 Specificity of User Queries

On occasion, users entered straightforward queries into the experimental online catalog, retrieved seemingly useful titles, but rated them "possibly useful" or "not useful." We have used the rubric "Specificity of user queries" to describe these queries. We think that users might have had a different (and probably related) topic in mind but, for whatever reason, they entered a query into the experimental online catalog that was not indicative of their information needs. The project director encountered comparable queries in a previous research project (Markey and Demeyer 1986, 157–8).

Table 10.3 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to specificity of user queries.

Table 10.3. Specificity of User Queries

User query	Explanation
computer user interface	Blue System retrieves one title in a keyword-in-record search. The user gives the title (*Designing the user interface: strategies for effective human-computer interaction*) a "not useful" rating.
composite materials	Pinstripe System retrieves one subject heading, i.e., "Composite materials," in a keyword-in-subdivided-heading search. User selects the heading, displays all 13 retrieved titles, and gives them negative relevance assessments. Examples of these titles are *Mechanics and composite materials, Composite materials*, and *Composite materials and their use in structures*.
material handling	Blue System informs the user that the entered term is a cross reference to "Materials handling." Exact-approach options for displaying general works, broader terms, narrower terms, subject and form subdivisions, and for expanding the search are enabled. The user displays several titles (e.g., *Integrated materials handling in manufacturing, The principles and practice of materials handling*) under subdivided forms of this subject heading and gives them negative relevance assessments.

Had we conducted post-search interviews with users, we would have had explanations from users regarding their reasons for giving negative relevance assessments to retrieved titles. In the absence of interview data, we can speculate that the topic that users had in mind was not quite the same as the subjects of the queries they entered into the experimental online catalog.

A total of seventeen searches failed because of query specificity. Twelve searches were conducted in the Blue System and five searches were conducted in the Pinstripe System. The failure analysis demonstrates that these searches were really successful searches in terms of leading users to titles on the topics they expressed in the queries they entered into the catalogs. In view of user relevance assessments, the searches were not successful in terms of retrieving titles that fulfilled the specifications of queries users had in their minds but did not make explicit in their queries.

10.2.1.3 Conflicting Relevance Assessments

Section 8.5.5 described how we resolved conflicting relevance assessments. Basically, we used the last relevance assessment that searchers gave to titles displayed more than once and disregarded other relevance assessments for the same title.

Conflicting relevance assessments could have affected a search in a single system. For example, a user could have given two different relevance assessments to the same title in a search of the Blue System or to the same title in a search of the Pinstripe System. Conflicting relevance assessments could have also affected searches of both experimental systems for the same topic. That is, a user could have given two different relevance assessments to the same title in separate searches of the Blue and Pinstripe Systems.

The researchers attributed the failure of eight searches (five in the Blue System and three in the Pinstripe System) to conflicting relevance assessments. An example of a search in which we attributed failure to conflicting relevance assessments was a search of the Blue System for "attenuation equalizers." The Blue System informed the user that this term was a cross reference to the subject heading "Equalizers (Electronics)" and responded with several exact-search options. The user displayed the title *Passive equalizer design data* three times, rated it "very useful" the first and second times, and "not useful" the third time.

10.2.1.4 Database Failure

The failure of searches was attributed to database failure in the following circumstances: (1) users retrieved titles but displayed titles were much more specific than the original user query, (2) users retrieved too few titles on topics that were covered in the selected classification areas, and (3) users retrieved no titles on topics that were covered in the selected classification areas. The researchers were guided in their decision about attributing failure to database failure on the titles users displayed in their searches and the negative relevance assessments they gave to titles.

Table 10.4 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to database failure.

Table 10.4. Database Failure

User query	Explanation
workstation	Blue System retrieves one title in a title-keyword search, *Workstation environment for wastewater treatment design using artificial intelligence and mathematical models,* which the user rates "not useful." In Pinstripe, the keyword-in-subdivided-heading search retrieves no headings.
virtual reality	Both Blue and Pinstripe Systems fail to retrieve titles. Blue System enlists subject searching approaches for queries composed of more than one word. Pinstripe System tries to retrieve titles through the keyword-in-record search.
design propeller	Pinstripe System invokes the alphabetical approach which places the user in an alphabetical list of subject headings at the subject headings "Design, Industrial" and "Desktop publishing." Blue System enlists subject searching approaches for queries composed of more than one word and fails to make a match.

Examples of other queries that failed to retrieve useful titles in both Blue and Pinstripe Systems due to database failure were:

- linked lists
- oriented polymers
- time motion studies
- multiple integration
- modem programming
- wide area networks
- joint application development

The failure of 120 searches was attributed to database failure. Of these 120 searches, there were 59 pairs of searches in which the failure of searches in both the Blue and Pinstripe Systems was attributed to database failure. The failure of another two searches was attributed to database failure in the Pinstripe System. The failure of Blue-system searches for the same topics was attributed to the exact approach because users seemed to have difficulty navigating exact-search options in the Blue System. An example was a Blue-system search in which the user was interested in historic preservation and began the search with the query "architecture." The Pinstripe System responded with 16 subject headings retrieved in a keyword-in-subdivided-heading search. The user displayed titles assigned to "Landscape architecture" and "Architecture — Conservation and restoration" and gave them negative relevance assessments. In the search of the Blue System for the same topic, the user seemed to tire of selecting the several exact-approach options that led to the display of the subject heading "Architecture — Michigan — Detroit — Conservation and restoration" because she terminated the search before displaying titles assigned to this subdivided subject heading. Section 10.2.1.8 focuses on navigational problems in experimental-system searches.

10.2.1.5 Large Retrievals

We did not expect large retrievals to be a problem in this study because the databases of the experimental online catalog were quite small. Yet, a few topics were represented by large numbers of bibliographic records, e.g., computers, architecture, various computer languages, mathematics.

Table 10.5 details one Blue-system search and three Pinstripe-system searches that were plagued by too many retrievals.

Table 10.5. Large Retrievals

User query	Explanation
fortran	Blue System responds with the results of an exact search in which options for displaying general works, narrower terms, form subdivisions, and for expanding the search are enabled. The user selects the general works option. The system responds by displaying the first of 110 retrieved titles. The user walks away from the system without even rating the first title.
computer	Pinstripe System responds by displaying the first of 2,324 titles retrieved in a keyword-in-record search. User displays three titles and rates them "not useful."
materials	Pinstripe System responds by displaying the first of 815 titles retrieved in a keyword-in-record search. User rates this title "not useful" and ends the search.
operations research	Pinstripe System retrieves 144 subject headings in a keyword-in-subdivided-heading search. The user selects one subject heading, displays one title, gives it a "not useful" rating, and ends the search.

Of the 14 searches that failed due to the retrieval of too many titles, 13 searches were conducted in the Pinstripe System. Most of the time, large retrievals in the Pinstripe System were connected with the results of keyword-in-record searches; keyword-in-subdivided-heading searches occasionally produced too many titles.

As library catalogs have increased in size in recent years, several researchers have focused on the problem of too many retrievals (Markey and Demeyer 1986, 277; Wiberley and Daugherty 1988; Lynch 1990). In a recent study, Wiberley, Daugherty, and Danowski (1993) concluded that few users displayed titles when the number of retrieved titles exceeded 200. They suggested that system designers provide features that help users cope with retrievals that exceed this number (Wiberley, Daugherty, and Danowski 1993, 5). In view of their research, we considered searches which resulted in large retrievals (titles or subject headings) as unsuccessful searches because the particular subject searching approach invoked by the experimental system did not provide features for managing the display of such large retrievals.

10.2.1.6 Vocabulary of User Queries

The experimental online catalog contained information on the topics that interested several users. However, users phrased their queries in certain ways that prevented the experimental online catalogs from making retrievals in response to their queries. When users entered queries with such phrasing, the two experimental systems responded in different ways. The Blue System exhausted the subject-searching approaches on the search trees before responding to users that it was unable to find titles to satisfy their queries. The Pinstripe System performed the search using one of the three subject-searching approaches at its disposal before responding to users that it was unable to find titles to satisfy their queries. However, differences between the two systems in terms of subject-searching approaches were transparent to end users.

An example of a search characterized with an unfortunate phrasing of the query was for "mechanical statics." The Blue System exhausted the subject-searching approaches on the search trees and informed the user that it was unable to find titles to satisfy the query. The Pinstripe System submitted this query to a keyword-in-subdivided-heading search and made no retrievals. If the systems had automatically truncated the word "mechanical," they would have retrieved titles like *Engineering mechanics statics*, *Schaum's outline of theory and probabilities of engineering: mechanics, statics, and dynamics*, and *Engineering mechanics: statics and dynamics*.

Table 10.6 describes four of the 31 queries that failed due to the vocabulary of user queries. In the left column of Table 10.6 are listed queries that failed due to vocabulary problems. Under the right-hand column are suggested ways in which systems could process queries to retrieve titles and subject headings on users' topics of interest. Fourteen queries had vocabulary problems that adversely affected both searches in Blue and Pinstripe systems. Vocabulary problems adversely affected another three queries in the Blue System only. (The search for the same topic in the Pinstripe System failed for different reasons.)

Table 10.6. Vocabulary Problems

User query	Explanation
flow through porous media	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If this query were reduced to "porous media," the experimental system would retrieve *Fundamentals of transport phenomena in porous media* in a title-keyword search.
serial communication	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If automatic truncation were applied to this query, the experimental system would retrieve *Mastering serial communications* in a title-keyword search.
factor of safety	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If this query were reduced to "safety," the experimental system would respond with the alphabetical approach and display subject headings such as "Safety engineering," "Safety, Industrial," "Safety of systems," and "Safety of workers."
c.a.s.e.	Blue System retrieves seven titles bearing long contents notes which contain various initials for author first and middle names including the letters and punctuation "c.," "a.," "s.," and "e." Pinstripe System's keyword-in-subdivided-heading approach fails to produce retrievals. The experimental system should delete punctuation in user queries. If single letters result, the system should concatenate the letters to form an acronym. If two or more series of characters result, the system should treat the query as a phrase and submit it to the search tree for queries exceeding one word.

If we could redesign the ASTUTE experimental online catalog, we would add system capabilities such as automatic stemming and/or best match to improve retrieval for the queries in this failure analysis category. We would also remove punctuation in user queries to effect matches of acronyms, subject headings, or title words. (Section 12.3 focuses on searches with vocabulary problems and makes suggestions for system improvements.)

10.2.1.7 Subject Searching Approach Failures

The experimental online catalog contained information on topics that interested users. However, its failure to retrieve the desired information was occasionally attributed to the particular subject searching approach selected by the system. For example, in response to the query "chemical engineering," the Pinstripe System performed a keyword-in-subdivided-heading search and failed retrieve subject headings. (The user's search in the Blue System for this query failed for a different reason, namely, perseverance.)

Nineteen subject searching approach failures occurred in Pinstripe-system searches. The Pinstripe System failed to produce potentially useful retrievals through the subject searching approach that it chose randomly. (The Blue-system searches for the same topics failed for different reasons, e.g., perseverance, conflicting relevance assessments, vocabulary problems.)

Table 10.7 lists user queries and gives explanations for subject searching approach failures in Pinstripe-system searches.

Table 10.7. Subject Searching Approach Failures in the Pinstripe System

User query	Explanation
control system	Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (A vocabulary problem adversely affects the Blue-system search for this query.)
bar code	Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (Search approach also adversely affects the Blue-system search for this query.) Had the systems responded with the alphabetical approach, they would have placed the user close to the cross reference "Bar coding of products."
attenuation equalizers	Pinstripe System responds to this query with the keyword-in-record approach and is unsuccessful retrieving titles. (Conflicting relevance assessments adversely affect the Blue-system search for this query that leads to the retrieval of titles through the cross reference "Attenuation equalizers" for the subject heading "Equalizers (Electronics).")
material handling	Pinstripe System responds to this query with the keyword-in-record approach and retrieves two titles that the user rates "possibly useful." (A navigational problem adversely affects the Blue-system search for this query which features the exact approach due to the cross reference "Material handling" for the subject heading "Materials handling.")

cad	Pinstripe System responds to this query with the keyword-in-subdivided-heading approach and is unsuccessful retrieving subdivided headings. (Perseverance adversely affects the Blue-system search for this query in which the alphabetical approach displays an alphabetical list of subject headings bearing the cross reference "CAD" for the subject heading "Computer-aided design." Instead of selecting the reference, the user ends the search.)

Of the 19 search approach failures in the Pinstripe System, all but two searches failed to produce retrievals through the keyword-in-subdivided-heading approach. The two remaining failures involved the keyword-in-record approach and were described in Table 10.7.

Search approach failures were not as common in the Blue System as in the Pinstripe System. Table 10.8 lists user queries and gives explanations for the three subject searching approach failures in Blue-system searches.

Table 10.8. Subject Searching Approach Failures in the Blue System

User query	Explanation
dataflow	Blue System exhausts the search trees and fails to make a match. (Perseverance adversely affects the Pinstripe-system search for this query in which the user needs to browse forward 2 screens to find cross references such as "Data flow across national boundaries" and "Data flow, Transborder" that describe the user's topic of interest.)
flow charting	Blue System exhausts the search trees and fails to make a match. (Pinstripe System responds to this query with the keyword-in-subdivided-heading approach and is unsuccessful retrieving subdivided headings.) Had the systems responded with the alphabetical approach, they would have placed the user close to the subject heading "Flow charts."
bar code	Blue System exhausts the search trees and fails to make a match. (Pinstripe System responds to this query with the keyword-in-subdivided-heading approach and is unsuccessful retrieving subdivided headings.) Had the systems responded with the alphabetical approach, they would have placed the user close to the cross reference "Bar coding of products" for the subject heading "Product coding."

In all three search approach failures, the Blue System exhausted the search trees without making a match that would have invoked a particular subject searching approach. Yet all three failures could have been successful searches had the Blue System responded with the alphabetical approach. Searchers would have still have to browse listed cross references or subject headings and select them to retrieve titles on their topics of interest.

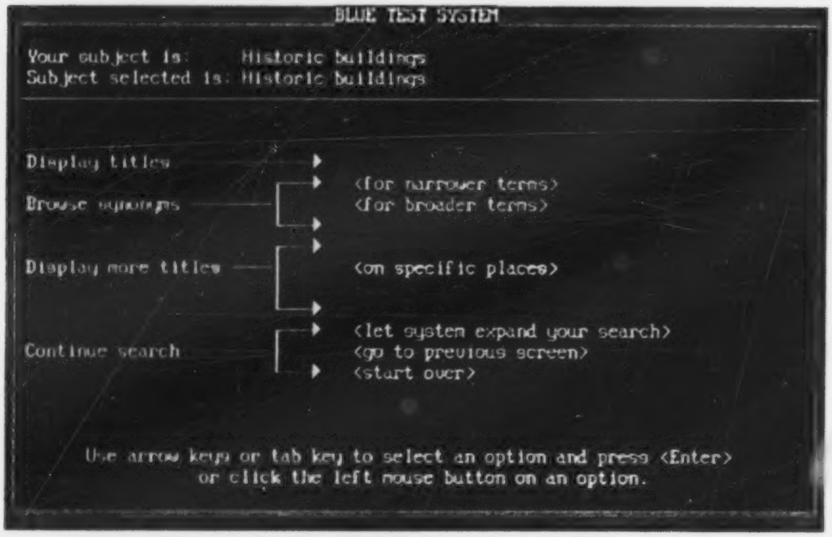
In searches in which systems exhaust search trees and fail to make retrievals, it is possible users would find a pertinent search term in lists of subject headings in the alphabetical neighborhood of the user-entered query. Thus, the alphabetical approach could be the recommended system response when systems exhaust search trees without finding a match.

10.2.1.8 Navigational Problems

Subject searching approaches that featured matches of controlled vocabulary terms (e.g., exact, alphabetical, keyword-in-main-heading) offered users many options for furthering their search. On occasion, users pursued several potentially fruitful options but, despite their searching, they were unable to find useful titles. We retraced their actions in an effort to increase our understanding of what they experienced during their search. In the course of recreating their searches, we realized that users might have experienced difficulties navigating the experimental online catalog. In addition to navigational problems, users could have become frustrated with how long the system took to search selected subject headings, broader or narrower terms, or subdivided forms of selected subject headings.

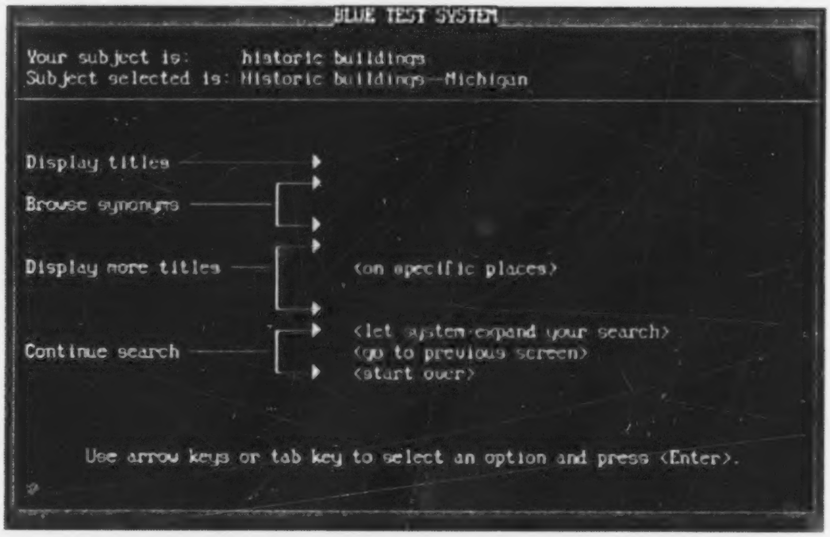
An example of a search plagued by navigational problems and user frustration at system response time was a search for historic preservation. The user initiated this search with the query "architecture." The Blue System responded with the results of an exact search in which options for displaying general works, broader terms, narrower terms, subject, place, and form subdivisions, and for expanding the search were enabled. The user browsed narrower terms and selected "Historic buildings." It took the Blue System about 30 seconds to search for the subject heading "Historic buildings," determine which exact-search options were enabled, and create the exact-search main menu (see figure 10.1).

The user selected the place-subdivision option and choose the subdivision "Michigan." It took the Blue System another 30 seconds to search for "Michigan" under the main subject heading "Historic buildings." Since there were subdivided forms of the subject heading "Historic buildings— Michigan," the system also had to determine which exact-search options were enabled and create the exact-search main menu for the subject heading "Historic buildings— Michigan" (see figure 10.2).



```
                         BLUE TEST SYSTEM
 Your subject is:      Historic buildings
 Subject selected is: Historic buildings

 Display titles ——————————————▶
                             ▶   <for narrower terms>
 Browse synonyms ————————┤       <for broader terms>
                             ▶
                             ▶
 Display more titles ————┤       <on specific places>
                             ▶
                             ▶   <let system expand your search>
 Continue search ————————┤       <go to previous screen>
                             ▶   <start over>

       Use arrow keys or tab key to select an option and press <Enter>
               or click the left mouse button on an option.
```

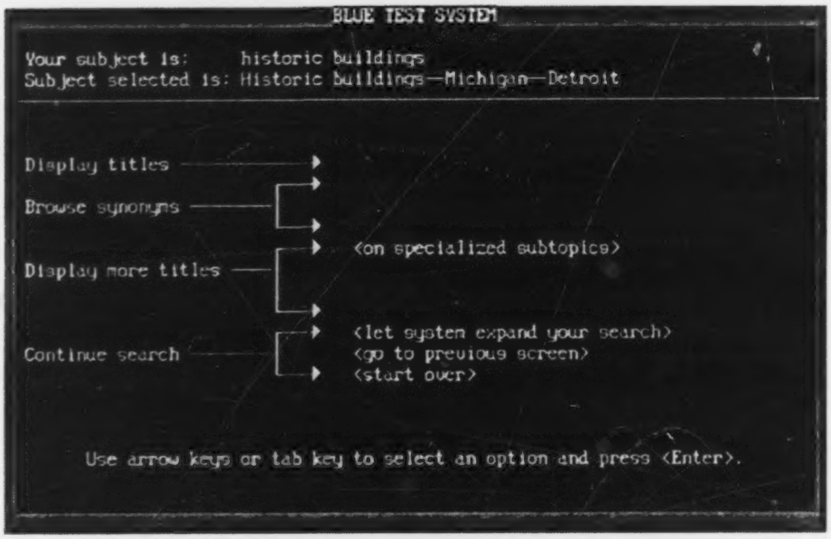
Figure 10.1. Exact-search options for "Historic buildings"



```
                         BLUE TEST SYSTEM
 Your subject is:      historic buildings
 Subject selected is: Historic buildings--Michigan

 Display titles ——————————————▶
                             ▶
 Browse synonyms ————————┤
                             ▶
                             ▶
 Display more titles ————┤       <on specific places>
                             ▶
                             ▶   <let system expand your search>
 Continue search ————————┤       <go to previous screen>
                             ▶   <start over>

       Use arrow keys or tab key to select an option and press <Enter>.
```

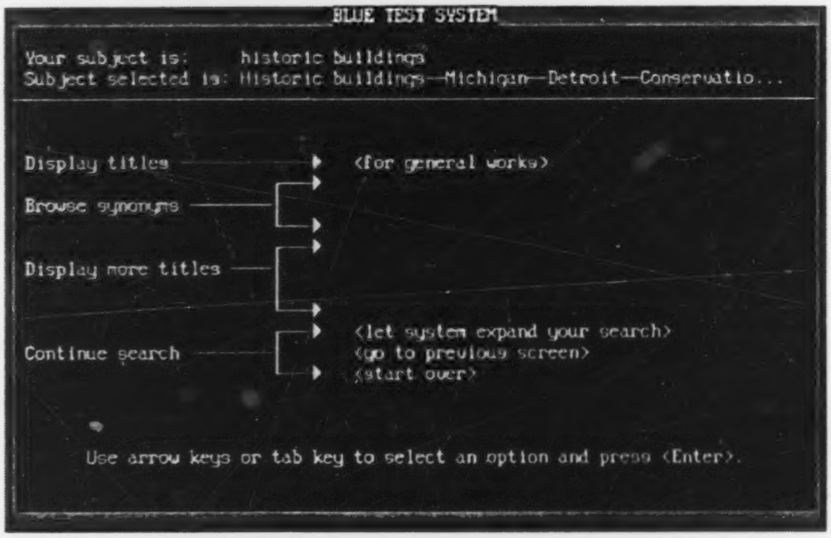
Figure 10.2. Exact-search options for "Historic buildings— Michigan"

BEST COPY AVAILABLE



```
                         BLUE TEST SYSTEM
Your subject is:     historic buildings
Subject selected is: Historic buildings—Michigan—Detroit

Display titles ──────────────►
Browse synonyms ─────────┬───►
                         └───►
                         ┌───►  <on specialized subtopics>
Display more titles ─────┤
                         └───►
                         ┌───►  <let system expand your search>
Continue search ─────────┤      <go to previous screen>
                         └───►  <start over>

      Use arrow keys or tab key to select an option and press <Enter>.
```

Figure 10.3. Exact-search options for "Historic buildings— Michigan — Detroit"



```
                         BLUE TEST SYSTEM
Your subject is:     historic buildings
Subject selected is: Historic buildings—Michigan—Detroit—Conservatio...

Display titles ──────────────►  <for general works>
Browse synonyms ─────────┬───►
                         └───►
                         ┌───►
Display more titles ─────┤
                         └───►
                         ┌───►  <let system expand your search>
Continue search ─────────┤      <go to previous screen>
                         └───►  <start over>

      Use arrow keys or tab key to select an option and press <Enter>.
```

Figure 10.4. Exact-search options for "Historic buildings— Michigan — Detroit — Conservation and Restoration"

The Blue System responded once again with the results of an exact search in which options for displaying place subdivisions and for expanding the search were enabled.

BEST COPY AVAILABLE

The user selected the place-subdivision option again and chose the subdivision "Detroit." It took the Blue System another 30 seconds to search for "Detroit" under the subdivided subject heading "Historic buildings— Michigan." The Blue System responded once again with the results of an exact search in which options for displaying subject subdivisions and for expanding the search were enabled (see figure 10.3).

The user selected the subject-subdivision option and chose the subdivision "Conservation and restoration." It took the Blue System another 30 seconds to search for "Conservation and restoration" under the subdivided subject heading "Historic buildings— Michigan — Conservation and restoration." The Blue System responded once again with the results of an exact search in which options for displaying general works and for expanding the search were enabled (see figure 10.4). At long last, the system presented the user with the opportunity to display a title. For some reason, the user did not select the general-works option and terminated the search.

In the absence of user interview data, we speculated on the reasons why the user terminated the search. Perhaps she was frustrated by the length of time it took to select an exact-search option and obtain a useful response. Perhaps she lost her bearings in this lengthy, complex search and preferred to start the search over again.

In Table 10.9, two more user queries are described. The failure of these two searches was attributed to navigational problems.

The failure of five Blue-system searches was attributed to navigational problems. Three of the five searches were described in this section. Scenarios of the two remaining failed searches were quite similar to the descriptions of the two searches given in Table 10.9.

The Blue System's exact search played a major role in the failure of these five searches. Several improvements are needed to reduce user navigational problems in operational systems. First, the response time for searching selected terms and creating exact-search option menus should be reduced. Second, exact-search option menus should be used to summarize large retrievals only. If the selection of an exact-search option in the original main menu retrieves a few subject headings or titles, the system should go directly to the titles display rather than inserting exact-search option menus or subject heading lists as intermediary steps. Third, when the system matches a cross reference, it should check to determine whether the "expand search" option produces additional matches and disable the option if it does not produce such matches.

Table 10.9. Navigational Problems

User query	Explanation
material handling	Blue System informs user that the entered term is a cross reference for "Materials handling." Blue System responds with enabled exact-search options for displaying general works, narrower terms, broader terms, subject and form subdivisions, and expanding the search. User chooses form-subdivisions option. System responds with enabled exact-search options for displaying general works and expanding the search. Instead of displaying general works, the user expands the search. The system exhausts the search trees and does not find another match. (The initial match was made on the cross reference "Material handling" for the subject heading "Materials handling.")
audio amplifier	Blue System informs user that the entered term is a cross reference for "Amplifiers, Audio." Blue System responds with enabled exact-search options for displaying subject subdivisions and expanding the search. User expands the search. Blue System responds with alphabetical list of subject headings. User chooses a cross reference for "Equalizers (Electronics)." User expands again and system performs a keyword-in-main-heading search. The result is the subject heading "Amplifiers, Audio" which the user selects. Despite the availability of a general-works option in the exact-approach main menu, the user does not display titles connected with this subject heading.

The ASTUTE experimental online catalog is the first system in which the exact approach has been implemented. This approach was far more successful in terms of delivering useful titles to users (see section 11.2) than it was unsuccessful. However, the failure analysis in this chapter provides an excellent opportunity to study the performance of this approach, identify problems, and suggest improvements.

10.2.1.9 Failure of Both Systems Revisited

Pairs of searches that failed to retrieve useful retrievals in both systems numbered 128 — a little over two-fifths of the pairs of searches performed in the experimental online catalog at UM-D. A failure analysis of these searches demonstrated the following problems that adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failures, (5) large retrievals, (6) user query vocabulary, (7) search approach failures, and (8) navigational problems.

Problems with user perseverance, specificity of user queries, and conflicting relevance assessments had little to do with the ability of systems to retrieve useful information. Sections 10.2.1.1 to 10.2.1.3 gave several examples of searches that would have retrieved potentially useful titles but for user actions that prevented retrieval. In this failure analysis, we considered such searches successful.

Database failures are bound to occur in searches of any database. And users are likely to search databases for information on topics that the database does not cover or covers only marginally. Database failures are a type of system failure. Additional research is needed to determine the characteristics of queries that are likely to fall victim to database failures so that systems can redirect such queries to database(s) that contain the desired material. Other failures connected with systems were due to large retrievals, user query vocabulary, subject searching approach failures, and navigational problems. Throughout the subsections that described searches that failed for these reasons, we offered suggestions for making improvements to systems to reduce failure in future experimental and operational systems.

Table 10.10 reclassifies unsuccessful searches in the Blue and Pinstripe Searches. The reclassification is limited to the 128 pairs of unsuccessful searches conducted in the Blue and Pinstripe Systems and featured in subsections of section 10.2.1.

Table 10.10. Reclassified Searches
(Both systems, UM-D)

	Blue System		**Pinstripe System**	
Useful titles retrieved	**No.**	**%**	**No.**	**%**
Successful searches†	43	33.6	21	16.4
Database failures	56	43.8	58	45.3
System failures*	29	22.6	49	38.3
Total	128	100.0	128	100.0

†Barring perseverance, query specificity, conflicting relevance assessments.

*Including problems with large retrievals, subject searching approaches, user query vocabulary, navigation.

Over forty percent of unsuccessful searches in the Blue and Pinstripe Systems were attributed to database failure. A larger percentage (38.3%) of system failures was connected with the Pinstripe System than with the Blue System (22.6%). Many system failures were search approach failures in the Pinstripe System. The failure analysis promoted some unsuccessful searches to successful searches. One-third of Blue-system searches and one-sixth of Pinstripe-system searches were successful in terms of guiding users to titles on the topics their queries described. Many of the searches promoted to successful status were due to perseverance in the Blue System.

Figure 10.5 is a graphical depiction of the statistics in Table 10.10. It shows obvious differences in the percentages of successful searches and system failures in Blue and Pinstripe Systems.

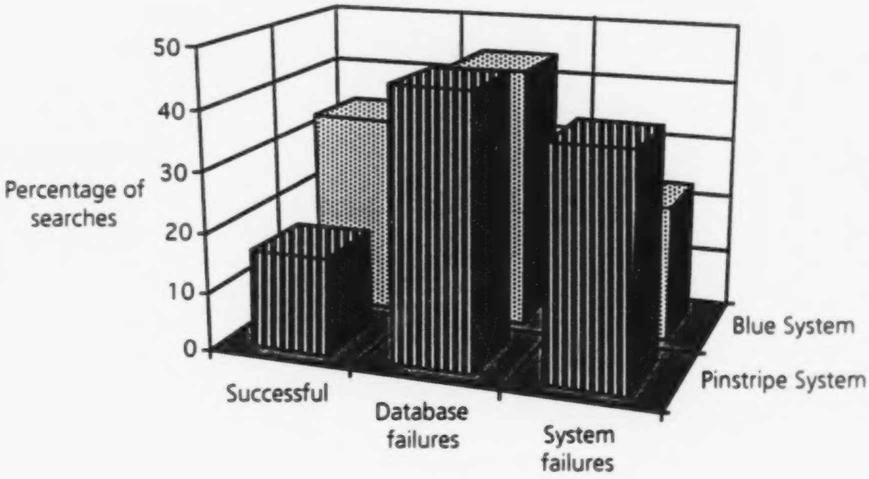


Figure 10.5. Reclassified unsuccessful searches (both systems, UM-D)

10.2.2 Failure of One System

A total of 147 searches failed to yield at least one useful title in one of the two experimental systems (see Table 10.1). Users performed 69 searches in the Blue System and 78 searches in the Pinstripe System. Of the total of 147 unsuccessful searches, 95 searches were successful in retrieving at least one useful title in one of the two experimental online catalogs. The failure analysis involved recreating the search in both systems — the system that failed to yield useful titles and the system that was successful in producing useful titles — to determine the reason for failure in the former systems. Users performed the remaining 52 (of 147) searches in only one of the two experimental online catalogs. The failure analysis involved recreating the search in the one system searched by the user. It also involved performing a search for the same topic in the other system to determine whether users would have been successful retrieving useful titles had they searched two systems for their topics of interest.

This failure analysis of searches resulted in a categorization of searches based on the problems that adversely affected the retrieval of potentially useful information from the experimental systems. Problems were the same as problems for the failure of searches in both systems except for the addition of a new category for spelling problems. The nine categories were as follows: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failures, (5) large retrievals, (6) user query vocabulary, (7) subject search approach failures, (8) navigational problems, and (9) spelling problems. Subsequent sections discuss these problems as they pertain to unsuccessful searches in one of the two experimental systems.

10.2.2.1 User Perseverance

User perseverance was a common reason for the failure of searches to retrieve useful titles. For example, a user entered the subject query "turbo prolog" into the Blue System. It responded with an exact-approach main menu with enabled options for displaying titles, broader terms, and expanding the search. The user ended the search. Had the user selected the titles-display option, he would have retrieved the title *Introduction to Turbo Prolog.* This and the many other searches like it should not be counted as unsuccessful searches. The catalog retrieved titles on the subject queries that users entered. Users chose to terminate their searches rather than display titles.

Perseverance was an especially frequent reason for the failure of searches in the Blue System. Perseverance was attributed to the failure of over two-thirds (68.1%) of the 69 failed searches in the Blue System. In contrast, perseverance was attributed to the failure of a little over one-third (37.2%) of the 78 failed searches in the Pinstripe System.

Table 10.11 recounts user and system activity that resulted in attributing the failure of three more searches conducted in the Blue System to user perseverance.

Two of the three user perseverance failures described in Table 10.11 were connected with controlled vocabulary search approaches in the Blue System. Of the 47 failures due to user perseverance in the Blue System, 38 failures involved controlled vocabulary search approaches. Subject searching approaches that featured matches of controlled vocabulary terms (e.g., exact, alphabetical, keyword-in-main-heading) offered users many options for furthering their search. Perhaps users were overwhelmed with the array of available options and made actions to terminate their searches or start them over in a different way. Failure due to perseverance was attributed to 29 searches in the Pinstripe System but there was not the preponderance of such failures for controlled vocabulary approaches in the Pinstripe System as there was in the Blue System.

Table 10.11. User Perseverance

User query	Explanation
flowcharting	Blue System retrieves 8 titles in a title-keyword search. The user walks away without rating even the first title.
computer aided design	Blue System responds with the results of an exact search in which options for displaying general works, narrower terms, subject and form subdivisions, and for expanding the search are enabled. The user ends the search.
auto	Blue System responds with the alphabetical approach bearing subject headings beginning with the word "auto," e.g., "Auto mechanics." User browses forward to "Automobile maintenance" which is a cross reference for "Automobiles — Motors — Maintenance and repair," selects it, but does not display retrieved titles such as *About your car: a basic guide to understanding and maintaining automobiles, Automotive diagnosis and tuneup,* and *The back-yard mechanic.*

Section 10.2.2.8 discusses navigational problems connected with controlled vocabulary search approaches in the Blue System and makes suggestions for improving the design of these approaches. These suggestions could also address problems associated with user perseverance, and, thus, reduce the numbers of users who walk away from the results of controlled vocabulary search approaches.

10.2.2.2 Specificity of User Queries

The problem of specificity of user queries adversely affected the retrieval of useful titles in four Blue-system searches and in four Pinstripe-system searches. The users whose searches were classified in this category probably had a different (and maybe related) topic in mind but, for whatever reason, they entered a query into the experimental online catalog that was not indicative of their information needs.

Table 10.12 recounts user and system activity that resulted in attributing the failure of two searches conducted in the Blue and Pinstripe Systems to specificity of user queries.

Table 10.12. Specificity of User Queries

User query	Explanation
environment	Blue System informs user that the entered term is a cross reference for "Ecology" or "Man — Influence of environment." User chooses the former. Blue System responds with enabled exact-search options for displaying general works, narrower terms, broader terms, notes, subject and form subdivisions, and expanding the search. User browses notes, narrower terms, and subject subdivisions. User displays one title assigned the narrower term "Ecology — Vocational guidance" and rates it "not useful." (In a Pinstripe-system search for this topic, the user rates 12 displayed titles "very useful" that were retrieved by the narrower term "Buildings — Environmental engineering.")
fluids	Pinstripe System places the user at "Fluids" in an alphabetical list of subject headings. User selects this subject heading and browses several broader terms and subject subdivisions. Of the total 25 titles displayed, the user rates none "very useful." (In Blue-system search for this topic, the user gives a "very useful" rating to 1 of 67 titles displayed through the exact approach. The title is *The analysis and design of pneumatic systems.*)

The descriptions of the two searches in Table 10.12 illustrated how users began with a query on a specific topic but the titles they rated "very useful" or the subject headings that delivered them to useful information were not always the same as their initial query.

10.2.2.3 Conflicting Relevance Assessments

The researchers attributed the failure of five searches (two in the Blue System and three in the Pinstripe System) to conflicting relevance assessments. An example of a search in which we attributed failure to conflicting relevance assessments was a search of the Pinstripe System for "vibration." The user began searching in the Blue System and retrieved 54 titles in a title-keyword search. She displayed nine titles and gave four titles "very useful" ratings. She searched the Pinstripe System for the same topic, retrieved 85 titles in a keyword-in-record search, and displayed the first of 85 titles. She gave this title a "not useful" rating and ended the search. This title was one of the four titles the user rated "very useful" in her Blue-system search for the same topic.

10.2.2.4 Database Failure

The failure of nine searches in the Blue System and the failure of nine searches in the Pinstripe System were attributed to database failure. Users performed eight of the nine Blue-system searches in the Blue System only. Users performed all nine Pinstripe-system searches in the Pinstripe System only. The one search performed in two systems was a search for "eiffel." The Blue System responded with the alphabetical approach and placed the user at the subject heading "Eiffel Tower (Paris, France)." The user ended the search. The user's Pinstripe-system search for the same topic yielded one title in a keyword-in-record search for the title *Gustave Eiffel* to which the user gave a "very useful" rating. Since the query immediately preceding this search had been given as "eiffel language," we felt that the user was interested in the "Eiffel" computer language and would have been more satisfied with a title we found in Michigan's Mirlyn online catalog — *Eiffel: the language*. Since the UM-D database did not contain this title, we considered this search of the Blue System to be a database failure.

Examples of other queries that failed to retrieve useful titles in both Blue and Pinstripe Systems due to database failure were:

- smalltalk
- propeller
- maximum shear stress criteria
- solar powered automobiles
- poka-yoke system

10.2.2.5 Large Retrievals

Large retrievals hampered six Pinstripe-system searches. Retrievals in all but one search were produced by keyword-in-record searches. Table 10.13 summarizes these searches.

Table 10.13. Large Retrievals

Query	Search approach	Retrievals
fortran	keyword-in-record	136 titles
science	keyword-in-record	1,207 titles
data structures	keyword-in-record	61 titles
materials	keyword-in-subdivided-heading	120 subject headings
basic	keyword-in-record	300 titles
architecture	keyword-in-record	106 titles

In all but one of the searches above, users walked away or they rated the first title that the system automatically displayed to them and then they walked away. In the search for "architecture," the user displayed 14 of 106 retrieved titles before terminating the search.

10.2.2.6 Vocabulary of User Queries

Vocabulary problems adversely affected three searches in the experimental online catalog. Users conducted one of the three searches in the Blue System and two of the three searches in the Pinstripe System. In the left column of Table 10.14 are listed queries that failed due to their vocabulary. Under the right-hand column are suggested ways in which systems could process queries to retrieve titles and subject headings on users' topics of interest.

Table 10.14. Vocabulary Problems

User query	Explanation
fundamental of data structures in c	Pinstripe System's keyword-in-subdivided-heading search fails to produce retrievals. If this query were reduced to "data structures in c," the experimental system would retrieve *Data structures: an advanced approach using C* in a title-keyword search.
how antilock brakes work	Pinstripe System's alphabetical approach places the user in an alphabetical list at the cross reference "How-to-do-it guides" for the subject heading "Do-it-yourself work." If this query were reduced to "antilock brakes," title-keyword and keyword-in-record searches would retrieve several titles.
reinforce concrete	Blue System exhausts search trees without retrieving titles or subject headings. If this query were truncated, the system would retrieve several titles bearing the subject heading "Reinforced concrete."

10.2.2.7 Subject Searching Approach Failures

Of the 24 search approach failures, only one failure occurred in the Blue System. This search began with the user-entered query "computers made simple." The Blue System exhausted the search trees without making a match. The Pinstripe System responded with the alphabetical approach. The user browsed forward in the alphabetical list to the subject heading "Computer," selected the heading, and gave "very useful" ratings to four titles. An example of a useful title was *High tech society: the story of information technology*. Table 10.15 summarizes search approach failures characteristic of searches in the Pinstripe System.

Table 10.15. Subject Searching Approach Failures in the Pinstripe System

User query	**Explanation**
car suspension	Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (Blue System retrieves 2 titles in a title-keyword search for this query that the user rates "very useful.")
hovercraft	Pinstripe System responds to this query with the keyword-in-record search and is unsuccessful retrieving titles. (In the Blue System, "hovercraft" is a cross reference for "Ground-effect systems." Exact approach yields titles such as *The glideway system: a high-speed ground transportation system in the northeastern corridor* ... and *Jane's surface skimmer systems*.)
tcp/ip	Pinstripe System responds to this query with the alphabetical approach and places the user at the subject heading "TI Programmable 59 (Calculator)." (Blue System retrieves 3 titles in a title-keyword search for this query that the user rates "very useful.")
cad	Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (In the Blue System, "cad" is a cross reference for "Computer-aided design" which leads the user to useful titles.)

In the first-listed search in Table 10.15, the Pinstripe System failed to retrieve titles through the keyword-in-subdivided-heading search and the Blue System was successful retrieving titles through a search selected by the search trees. Of the 23 Pinstripe-system searches that failed due to subject search approaches, this was by far the most common type of failure as it adversely affected 13 searches in the Pinstripe System.

Four searches failed because the Pinstripe System failed to retrieve titles through the keyword-in-record search. Blue-system searches for the same topic succeeded because the exact or alphabetical approaches in the Blue System featured cross references with terminology that matched user queries and delivered users to subject headings on their topics of interest. Pinstripe-system searches for these topics failed because the terminology used in cross references was not used in any subject-bearing fields of bibliographic records. Another four searches failed because the Pinstripe System responded with the alphabetical approach in which listed subject headings were not pertinent to users' topics of interest. Two Pinstripe-system searches failed because the subject heading(s) the system retrieved in a keyword-in-subdivided-heading search were not pertinent to users' topics of interest.

Subject searching approach failures in the Pinstripe System accounted for almost three-tenths (23 of 78 or 29.5%) of search failures. The Pinstripe System selected subject searching approaches randomly. Subject searching failures were rare in the Blue System, accounting for only one failure in 69 searches. The Blue System was guided in its selection of a subject searching approach by the search trees.

10.2.2.8 Navigational Problems

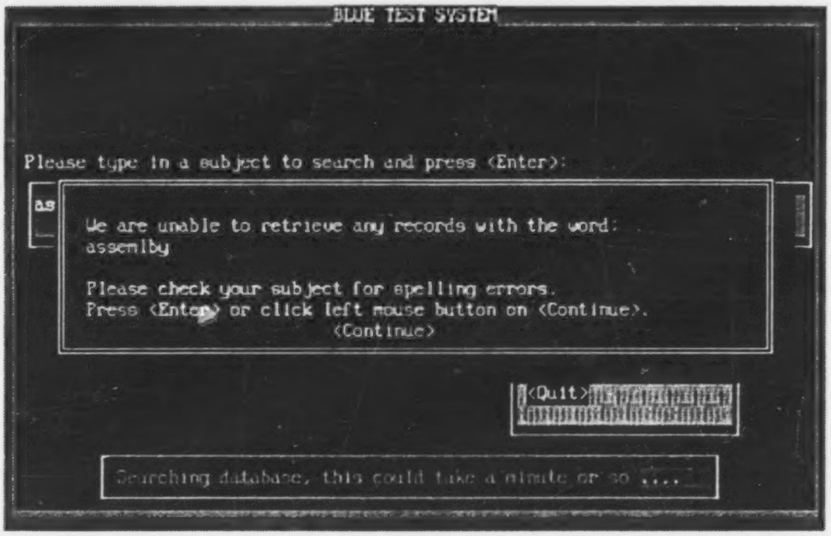
Navigational problems occurred in Blue-system searches. The two searches plagued by such problems required users to navigate the many options connected with controlled vocabulary approaches to find subject headings of interest. In a search for "animation," the Blue System responded to the user query with the alphabetical approach and placed the user near subject headings such as "Animation, Computer" and "Animated cartoons by computer." The user browsed forward to "Casting, Steel," selected this subject heading, and displayed titles even though it was not related to his original query "animation."

10.2.2.9 Spelling Problems

Spelling problems adversely affected the success of five searches in the experimental online catalog. Three searches were conducted in the Blue System and two searches were conducted in the Pinstripe System. In all five searches, users ignored the system's message that asked them to check their spelling and repeatedly clicked on the <Do search> option following the entry of their misspelled queries. These queries were:

- elecctronics
- assemlby language
- numeric rec
- software configration
- graphtheory

The systems specifically brought the potentially misspelled word to the user's attention. For example, figure 10.6 shows the system's response to the query "assemlby language" in which the system informed the user that the word "assemlby" might be misspelled.



```
BLUE TEST SYSTEM

Please type in a subject to search and press <Enter>:
as
   We are unable to retrieve any records with the word:
   assemlby

   Please check your subject for spelling errors.
   Press <Enter> or click left mouse button on <Continue>.
                         <Continue>

                                        <Quit>

  Searching database, this could take a minute or so ....
```

Figure 10.6. System response to misspelled query words

Computer-assisted spelling correction could have corrected these five queries. When users repeatedly ignore system messages to correct spelling, systems could automatically call up spelling correction algorithms and suggest corrections to possible spelling errors.

10.2.2.10 Failure of One System Revisited

Searches that failed to retrieve useful retrievals in one system numbered 147. A failure analysis of these searches demonstrated the following problems that adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failures, (5) large retrievals, (6) user query vocabulary, (7) search approach failures, (8) navigational problems, and (9) spelling problems.

Problems with user perseverance, specificity of user queries, and conflicting relevance assessments had little to do with the ability of systems to retrieve useful information. Sections 10.2.2.1 to 10.2.2.3 gave several examples of searches that would have retrieved potentially useful titles but for user actions that prevented retrieval. In this failure analysis, we considered such searches successful.

Database failures are bound to occur in searches of any database. And users are likely to search databases for information on topics that the database does not cover or covers only marginally. Database failures are a type of system failure. Other failures

BEST COPY AVAILABLE

connected with systems were due to large retrievals, user query vocabulary, subject searching approach failures, and navigational problems. Throughout the sections that described searches that failed for these reasons, we offered suggestions for making improvements to systems to reduce failure in future experimental and operational systems.

Table 10.16 reclassifies unsuccessful searches in the Blue and Pinstripe Searches. The reclassification is limited to the 147 unsuccessful searches conducted in both Blue and Pinstripe Systems and featured in section 10.2.2.

Table 10.16. Reclassified Searches (Single system, UM-D)

	Blue System		Pinstripe System	
Useful titles retrieved	**No.**	**%**	**No.**	**%**
Successful searches†	53	76.8	36	46.2
Database failures	9	13.0	9	11.5
System failures*	7	10.2	33	42.3
Total	69	100.0	78	100.0

†Barring perseverance, query specificity, conflicting relevance assessments.

*Including problems with large retrievals, subject searching approaches, user query vocabulary, navigation, spelling.

Database failure accounted for less than fifteen percent of unsuccessful searches in the Blue and Pinstripe Systems. A much larger percentage (42.3%) of system failures was connected with the Pinstripe System than with the Blue System (10.2%). Many of these failures were search approach failures in the Pinstripe System. The failure analysis promoted some unsuccessful searches to successful searches. Over three-quarters of Blue-system searches and a little under one-half of Pinstripe-system searches were successful in terms of guiding users to titles on the topics their queries described. Many of the searches were promoted to successful status were due to perseverance in the Blue System.

Figure 10.7 is a graphical depiction of the statistics in Table 10.16. It shows obvious differences in the percentages of successful searches and system failures in Blue and Pinstripe Systems.

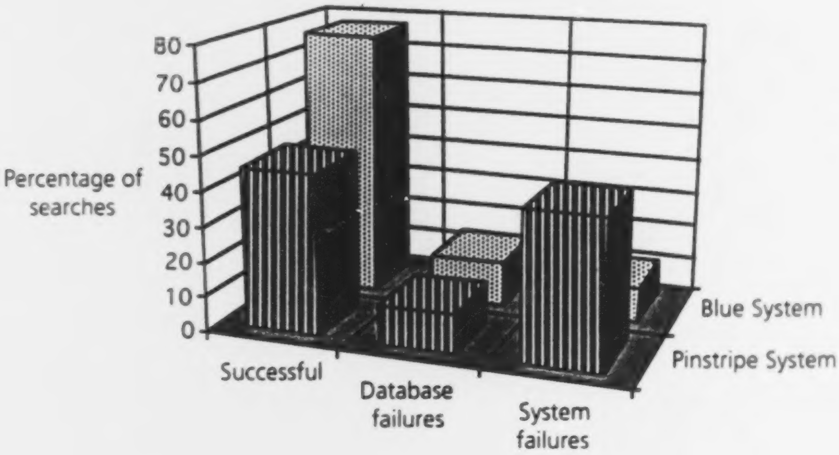


Figure 10.7. Reclassified searches (Single system, UM-D)

10.2.3 Revisiting Successful and Failed Searches

In the failure analysis of searches conducted by UM-D users, the researchers repeated user searches that failed to retrieve useful titles in an attempt to determine why they were unsuccessful in retrieving useful titles. In the process of repeating these searches, the failure analysis demonstrated that several searches might have been headed toward success but for various user actions that precluded the display of useful titles, e.g., perseverance, queries that were not indicative of user interests, conflicting relevance assessments. The failure analysis also demonstrated reasons for failure, e.g., design problems with a particular search approach or indexing technique, inability of a randomly-selected approach in the Pinstripe System to retrieve any titles in response to user queries. We can learn from the failure analysis in terms of designing subject searching approaches and reconfiguring the search trees. Chapter 12 makes recommendations for improving subject access in online catalogs based on failure analysis results.

Table 10.17 recasts the number and percentage of successful and unsuccessful searches based on the failure analysis of end-user searches at the University of Michigan-Dearborn.

Table 10.17. Reclassified Successful and Unsuccessful Searches at UM-D

Type of search	Blue System		Pinstripe System	
	No.	%	No.	%
Successful searches:				
≥1 useful titles retrieved	156	44.2	150	42.1
Retrieval of useful titles marred by perseverance	73	20.7	42	11.8
Retrieval of useful titles marred by query specificity	16	4.5	9	2.5
Retrieval of useful titles marred by conflicting relevance assessments	7	2.0	6	1.7
Subtotal	252	71.4	207	58.1
Unsuccessful searches:				
Database failure	68	19.2	70	19.6
Search approach failure	4	1.1	42	11.9
Large retrievals	1	0.3	19	5.3
Vocabulary of user queries	18	5.1	16	4.5
Navigation	7	2.0	0	0.0
Spelling	3	0.9	2	0.6
Subtotal	101	28.6	149	41.9
Total	353	100.0	356	100.0

A little under three-quarters of Blue-system searches were successful; about three-fifths of Pinstripe-system searches were successful. A large percentage (20.7%) of Blue-system searches would have been successful but users did not exhibit sufficient perseverance to stay with the search and see it to its conclusion. This percentage was somewhat smaller in the Pinstripe System (11.8%).

Although we have no user interview data, we can speculate on the reasons why many Blue-system users exhibited low levels of perseverance and terminated their searches prematurely. Our speculation is fueled by our experience searching the Blue and Pinstripe Systems and searching computer-based systems generally. We are quite sure that user perseverance in the Blue System was connected with amount of time the system required for users to browse the many options available in alphabetical and exact searches. These approaches allowed users to select listed subject headings, narrower terms, broader terms, or subdivided forms of selected subject headings. It took time for users to navigate between options and for the system to search for selected subject headings. Recommendations for improving alphabetical and exact searches made in the lengthy section on navigational problems included (section 10.2.1.8):

- Reducing the response time for searching selected terms and creating exact-search option menus.
- Enlisting exact-search option menus to summarize large retrievals.
- Requiring the system to check whether keyword searches on user-entered queries that match cross references retrieve additional titles and enable/disable the "expand search" option based on this system check.

Of the several reasons for unsuccessful searches, database failure affected about the same percentage (a little under twenty percent) of searches conducted in the Blue and Pinstripe Systems. Vocabulary of user queries was also troublesome and led to about five percent of search failures in both Blue and Pinstripe Systems. Large retrievals were usually connected with searches in the Pinstripe System. Disregarding database failures, the largest percentage of unsuccessful searches was for search approach failures in the Pinstripe System (11.9%). Only 1.1% of searches in the Blue System failed because of subject searching approaches. Much of the time, the Pinstripe System's keyword-in-subdivided-heading search was unable to retrieve subject headings bearing words in user queries or subject headings that were pertinent to users' topics of interest. On occasion, its alphabetical approach placed users in an alphabetical neighborhood that was not fruitful for enabling users to find satisfactory subject headings for expressing their topics of interest.

The Pinstripe System's selection of a subject searching approach in response to a user-entered query was based on a random selection algorithm. In contrast, the Blue System's selection of a subject searching approach was governed by search trees. These data provide evidence to support the development of search trees in online catalogs. In a catalog without search trees, a large percentage of end-user searches are bound to fail because systems do not assist end users in subject searching approach selection.

10.3 Successful and Unsuccessful Searches at Earlham

Section 10.3 gives the results of the failure analysis of searches performed in the Blue and Pinstripe Systems at Earlham College. Problems were the same as problems for the failure of searches in both systems and in a single system at UM-D. Recommendations for system improvements that were discussed in section 10.2 for unsuccessful searches at UM-D were also applicable to unsuccessful searches at Earlham.

Table 10.18 lists the number and percentages of successful and unsuccessful Blue and Pinstripe Searches.

Table 10.18. Successful and Unsuccessful Searches at Earlham

	Blue System		Pinstripe System	
Useful titles retrieved	**No.**	**%**	**No.**	**%**
≥1 useful titles retrieved	46	44.2	37	37.8
No useful titles retrieved in one of two systems	22	21.2	25	25.5
No useful titles retrieved in searches of both systems	36	34.6	36	36.7
Total	104	100.0	98	100.0

One or more useful titles were retrieved in 44.2% of searches in the Blue System and in 37.8% of searches in the Pinstripe System. The Blue System had a lower percentage (21.2%) of searches in which no useful titles were retrieved than in the Pinstripe System (25.5%). Thus, the figures in Table 10.18 indicate that the Blue System performed a little better than the Pinstripe System in terms of retrieving useful titles in searches. These percentages were quite similar to the percentages of successful and unsuccessful searches in the experimental online catalog at UM-D (see Table 10.1).

10.3.1 Failure of Both Systems

At Earlham, users performed a total of 86 pairs of searches. That is, they performed one search for their topic of interest in the Blue System followed by a search for the same topic in the Pinstripe System. Of the 86 pairs, 36 search pairs failed to yield useful titles in searches of both experimental systems — about 42% of the search pairs. The project director, an expert searcher of information retrieval systems, repeated these three dozen searches in both Blue and Pinstripe Systems to determine why end users were unsuccessful in retrieving useful titles.

10.3.1.1 User Perseverance

User perseverance was a common reason for the failure of searches to retrieve useful titles. For example, a user entered the personal-name query into the Blue System "jefferson" (last name), "thomas" (first name), and "inventions" (topic). The Blue System was unable to make retrievals through the keyword-in-record search for the combination of name and topic elements so it ignored the topic element and showed the user the results of an alphabetical search. That is, it placed the user in an alphabetical list of subject headings for personal names and highlighted the best matching subject heading, "Jefferson, Thomas, 1754–1826." The user selected this subject heading and displayed only the first of 62 retrieved titles. Had the user continued, the Blue System would have displayed titles on the user's topic of interest, e.g., *Thomas Jefferson: statesman of science, Thomas Jefferson: landscape architect, Jefferson and the arts,* and *Thomas Jefferson: the man, his world, his influence.*

Table 10.19 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to user perseverance.

Table 10.19. User Perseverance

User query	Explanation
women	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, notes, topical and place subdivisions, and for expanding the search are enabled. The user ends the search.
reconstruction	In the course of performing an alphabetical search, Pinstripe System gives the user a choice of two "Reconstruction" subject headings. The user selects the heading assigned to titles on the Reconstruction following the Civil War but he does not display titles connected with the selected heading.
reconstruction	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, notes, topical, place, and form subdivisions, and for expanding the search are enabled. The user browses and selects several subdivisions connected to the main heading but he does not display titles connected with the selected heading.
hoover herbert economic	Failing to produce retrievals for combinations of name and topic elements, Blue System ignores the topic element, responds with the alphabetical approach which highlights the personal-name heading "Hoover, Herbert, 1874–1964." User selects this heading and displays one of 28 retrieved titles. Had the user continued displaying titles, he would have found titles such as *The shattered dream: Herbert Hoover and the great Depression* and *The poverty of abundance: Hoover, the nation, the depression.*

If users had displayed several titles for the searches described in Table 10.19, they would have most likely succeeded in retrieving useful titles.

In the absence of interview data, we can speculate that the results of alphabetical or exact searches could have been too complex for users to understand. That is, they could have been lost or confused about which of the several options to pursue for continuing the search. This does not explain what was in the mind of the user who conducted the search for Herbert Hoover. The Blue System retrieved a manageable number of titles (28) and the user simply did not display additional titles.

A total of ten searches failed because of user perseverance. All but one of these searches were conducted in the Blue System. All nine searches conducted in the Blue System involved controlled vocabulary approaches. Perseverance was also a problem with failed searches at UM-D (see sections 10.2.1.1 and 10.2.2.1).

The failure analysis demonstrated that searches involving perseverance were really successful searches in terms of leading users to titles on their topics of interest. The problem was that users failed to display any or enough titles to find useful ones.

10.3.1.2 Specificity of User Queries

The problem of specificity of user queries adversely affected the retrieval of useful titles in four Blue-system searches and in three Pinstripe-system searches. Users whose searches were classified in this category probably had a different (and maybe related) topic in mind but, for whatever reason, they entered queries into the experimental online catalog that were not indicative of their information needs.

Table 10.20 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to specificity of user queries.

Perhaps, the topics that users had in mind were not quite the same as the queries they entered into the experimental online catalog. These searches were really successful searches in terms of leading users to titles on the topics they expressed in the queries they entered into the catalogs. However, the searches were not successful in terms of retrieving titles that fulfilled the specifications of queries users had in their minds but did not make explicit in their queries.

Table 10.20. Specificity of User Queries

User query	Explanation
boston tea party	Blue System responds with the results of an alphabetical search and highlights the subject heading "Boston Tea Party, 1773." The user selects the highlighted heading, displays the one title assigned this heading, i.e., *The Boston Tea Party*, and gives it a "not useful" rating.
richmond, indiana	Pinstripe System retrieves 16 titles through a keyword-in-record search. The user does not give "very useful" ratings to any of the 16 titles.
quaker	Blue System responds with the results of an alphabetical search that includes the subject headings "Quakers," "Quaker meeting houses," "Quaker church buildings." User displays titles assigned to these and other listed subject headings, expands the search by retrieving the results of a title-keyword search for this query. Examples of retrieved titles are *The Quaker colonies: a chronicle of the proprietors of the Delaware* and *The Dutch and Quaker colonies in America*. User does not give "very useful" ratings to the dozen titles displayed in this search.

10.3.1.3 Database Failure

Table 10.21 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to database failure.

In all three searches described in Table 10.21, the systems retrieved very few titles on the topics of user queries and displayed titles were probably too specific for user needs (e.g., the title on the hanging of a sheriff in nineteenth-century Montana in response to a query on 19th century Montana), or could have covered topics besides the one(s) users had in mind (e.g., the title on the witchcraft of native American peoples in response to the query on witchcraft). Post-search personal interviews would have provided explanations from users as to their reasons for giving negative relevance assessments to retrieved titles.

A total of 28 searches failed because of database failure. These 28 searches were pairs of searches, that is, searches conducted by fourteen different end users in the Blue and Pinstripe Systems.

Table 10.21. Database Failure

User query	Explanation
folk songs	Blue System responds with four subject headings retrieved in a keyword-in-subdivided-heading search. The user selects the subject heading "Folk songs, American — New York," displays one of two titles assigned this heading, i.e., *Body, boots, and britches: folktales, ballads, and speeches from country New York* and gives it a "not useful" rating.
witchcraft	Pinstripe System responds with three subject headings in a keyword-in-subdivided-heading search. The user selects the subject heading "Witchcraft — New England" and gives all three displayed titles negative relevance assessments. Examples of displayed titles are *Witchcraft and sorcery of the American native peoples* and *Navaho witchcraft*.
montana 19th century	Pinstripe System displays the one title *Hanging the sheriff: a biography of Henry Plummer* retrieved in a keyword-in-record search. User gives it a "not useful" rating.

A few searches in these pairs deserve special mention. Three searches were for combinations of names and topics:

- "king" (last name), "martin luther" (first name), "about the bycotts [sic] they had way back in mississippi" (topic)
- "king" (last name), "coretta scott" (first name), "learn about what happened in 1927" (topic)
- "king" (last name), "martin luther" (first name), "about in the 1925" (topic)

Since these three queries were alike in form, content, and were entered successively into the experimental online catalog, they were probably entered by the same individual. The Blue System responded to all three queries by ignoring the topic elements and placing users in an alphabetical list of personal-name headings where they could find the names "King, Coretta Scott, 1927–" and "King, Martin Luther, Jr., 1929–1968." The user selected these subject headings, displayed titles such as *Coretta: the story of Mrs. Martin Luther King, Jr.*, *There is a balm in Gilead: the cultural roots of Martin Luther King, Jr.*, and *Martin Luther King, Jr.: the making of a mind.* The user did not give positive relevance assessments to any displayed titles despite their promising content. Most likely, this user would have only given positive relevance assessments to titles that explicitly referred to the specific topic the user made explicit in his query.

The Pinstripe System performed keyword-in-record searches for all three queries and was unsuccessful retrieving titles for them. One could consider this a system failure instead of a database failure because the Pinstripe System failed to guide the user to potentially useful titles. However, if we attributed the failure of one of two searches in a pair to database failure, we almost always considered the second of the pair as a failure due to database failure. Other examples of queries that failed to retrieve useful titles in both Blue and Pinstripe Systems due to database failure are:

- dayton history
- women in reconstruction
- repeal of prohibition
- minorities in the west
- chicago strike

10.3.1.4 Large Retrievals

We did not expect large retrievals to be a problem in searches at Earlham College because the database of the experimental online catalog was quite small. Yet, a few topics were represented by large numbers of bibliographic records, e.g., Afro-Americans, Civil War, history, Indians, Jews, presidents, and slavery.

Large retrievals hampered two Pinstripe-system searches. In both searches, the system responded to the user-entered query "civil war" with a keyword-in-record search and retrieved 446 titles. In both searches, users failed to display more than the first of 446 retrieved titles.

10.3.1.5 Vocabulary of User Queries

Vocabulary problems adversely affected retrieval in 17 searches in the experimental online catalog. In the left column of Table 10.22 are listed four queries with vocabulary problems. Under the right-hand column are suggested ways in systems could process queries to retrieve titles and subject headings on users' topics of interest.

Of the seventeen queries that failed due to the vocabulary of user queries, eight queries had vocabulary problems that adversely affected both searches in Blue and Pinstripe systems. Vocabulary problems adversely affected one query in the Blue System; the Pinstripe-system search for this same query failed because of search approach problems. Thus, nine searches conducted in the Blue System and eight searches conducted in the Pinstripe System failed because of the vocabulary of user queries.

Table 10.22. Vocabulary Problems

User query	Explanation
black civil rights in the reconstruction	In the Pinstripe System, the query "reconstruction and civil rights" retrieves three titles in a keyword-in-record search: *Death at Cross Plains: an Alabama Reconstruction tragedy, Reconstruction, the Negro, and the New South*, and *Politics, principles, and prejudice, 1865–1866 — dilemma of Reconstruction America.*
hopi navajo indian land dispute	In the Blue System, query for "hopi navajo" retrieves one title in a keyword-in-record search, *The second long walk: the Navajo-Hopi land dispute.*
navajo crafts	In the Blue System, query for "navajo" results in an alphabetical search. Selection of the longer subject heading "Navajo Indians" results in an exact search options for subject subdivisions. Under the subdivisions "Material culture" and "Textile industry and fabrics" are potentially useful titles such as *Navajo material culture, Walk in beauty: the Navajo and their blankets*, and *Navaho weaving: its technic and its history.*
origins of the ku klux klan	In the Blue System, query for "ku klux klan" prompts the system to respond with the exact approach. The general works option produces titles such as *White terror: the Ku Klux Klan conspiracy and Southern Reconstruction* and *The Ku Klux Klan: a century of infamy*

In a large database, the vocabulary of some queries — "navajo crafts," "origins of the ku klux klan," "black civil rights during the reconstruction" — would not adversely affect the search because there would be a much larger database of unique words to combine in keyword searches. However, the vocabulary of the queries in this category of failed searches adversely affected retrieval in a small database such as the experimental online catalog database. If we could redesign the ASTUTE experimental online catalog, we would add system capabilities such as automatic stemming and/or best match to improve retrieval for the queries in this failure analysis category.

10.3.1.6 Subject Searching Approach Failures

The experimental online catalog contained information on topics that interested several users; however, it failed to retrieve the desired information. For example, in section 10.3.1.1, we discussed a search for "jefferson" (last name), "thomas" (first name), and "inventions" (topic) that failed to produce useful retrievals in the Blue System because of end-user perseverance. The user's search in the Pinstripe System failed for a different reason. That is, the Pinstripe System conducted a keyword-in-record search for this query and failed to retrieve titles.

The failure of eight searches was attributed to subject searching approach failures. All eight searches were conducted in the Pinstripe System. This system failed to produce potentially useful retrievals through the subject searching approach that the system chose randomly. Blue-system searches for the same topics failed for reasons connected with user perseverance, query specificity, or vocabulary.

Table 10.23 lists user queries and gives explanations for subject searching approach failures in Pinstripe-system searches.

Table 10.23. Subject Searching Approach Failures in the Pinstripe System

User query	Explanation
african-americans	The Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (Perseverance adversely affects the Blue-system search for this query.)
hoover herbert economic	The Pinstripe System responds to this query with the keyword-in-record search and is unsuccessful retrieving titles. (Perseverance adversely affects the Blue-system search for this query.)
slavery women	The Pinstripe System responds to this query with the keyword-in-subdivided-heading search and is unsuccessful retrieving subdivided headings. (Query specificity adversely affects the Blue-system search for this query.)
treatment of women slaves during civil war	The Pinstripe System responds to this query with the alphabetical approach. The user selects "Treaty of Ghent" but does not display titles assigned the selected heading. (Vocabulary adversely affects the Blue-system search for this query.)
women	The Pinstripe System responds to this query with the alphabetical approach. The user selects "Women spies," browses broader terms, selects "Spies," displays one title and deems it "not useful." (Perseverance adversely affects the Blue-system search for this query.)

Most subject searching approach failures were attributed to searches that completely failed to retrieve titles.

A few searching approach failures were connected with user problems navigating alphabetical lists of subject headings. For example, the two users who entered the queries "treatment of women slaves during civil war" and "women" in Table 10.23 seemed to have difficulty navigating the alphabetical list because the subject headings they selected had nothing to do with their original queries.

10.3.1.7 Failure of Both Systems Revisited

Pairs of searches that failed to retrieve useful retrievals in both systems numbered 36 — a little over two-fifths of the pairs of searches performed in the experimental online catalog at Earlham. A failure analysis of these searches demonstrated the following problems adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) database failures, (4) large retrievals, (5) user query vocabulary, and (6) search approach failures.

Problems with user perseverance and the specificity of user queries had little to do with the ability of systems to retrieve useful information. Sections 10.3.1.1 and 10.3.1.2 gave several examples of searches that would have retrieved potentially useful titles but for user actions that prevented retrieval. In this failure analysis, we considered such searches successful.

Database failures were considered a type of system failure. Other failures connected with systems were due to large retrievals, user query vocabulary, and subject searching approach failures; improvements can be made to systems to help them cope with these problems.

Table 10.24 reclassifies unsuccessful searches in the Blue and Pinstripe Searches. The reclassification was limited to the 36 pairs of unsuccessful searches conducted in both Blue and Pinstripe Systems and featured in subsections of section 10.3.1.

Table 10.24. Reclassified Searches (Both systems, Earlham)

	Blue System		Pinstripe System	
Useful titles retrieved	**No.**	**%**	**No.**	**%**
Successful searches†	13	36.1	4	11.1
Database failures	14	38.9	14	38.9
System failures*	9	25.0	18	50.0
Total	36	100.0	36	100.0

†Barring perseverance, query specificity.

*Including problems with large retrievals, subject searching approaches, user query vocabulary.

Almost forty percent of unsuccessful searches in the Blue and Pinstripe Systems were attributed to database failure. One-quarter of Blue-system searches and one-half of Pinstripe-system searches were due to system failures. The failure analysis promoted some unsuccessful searches to successful searches. A little over a third of Blue-system searches and a little over a tenth of Pinstripe-system searches were successful in terms of guiding users to titles on the topics their queries described.

Figure 10.8 is a graphical depiction of the statistics in Table 10.24. It shows obvious differences in the percentages of successful searches and system failures in Blue and Pinstripe Systems.

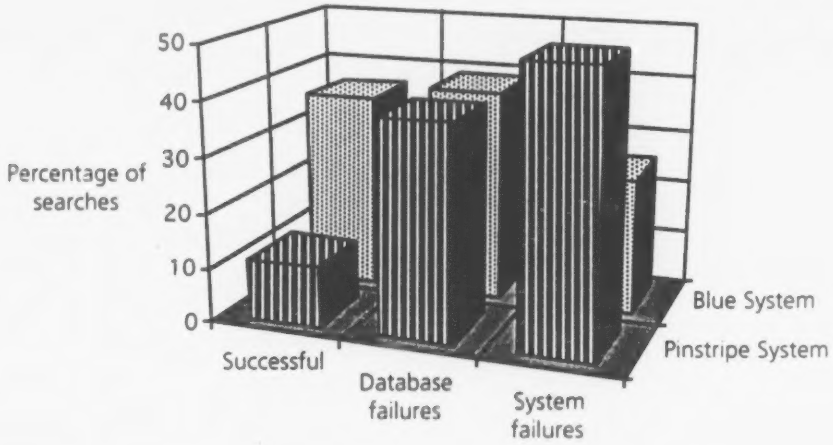


Figure 10.8. Reclassified searches both systems, Earlham)

10.3.2 Failure of One System

A total of 47 searches failed to yield at least one useful title in one of the two experimental systems (see Table 10.18). Users performed 22 searches in the Blue System and 25 searches in the Pinstripe System. Of the total of 47 unsuccessful searches, 35 searches were successful in retrieving at least one useful title in one of the two experimental online catalogs. The failure analysis involved recreating the search in both systems — the system that failed to yield useful titles and the system that was successful in producing useful titles — to determine the reason for failure in the former systems. Users performed the remaining 12 (of 47) searches in only one of the two experimental online catalogs. The failure analysis involved recreating the search in the one system searched by the user. It also involved performing a search for the same topic in the other system to determine the whether users would have been successful retrieving useful titles had they conducted their searches in two systems.

This failure analysis of searches resulted in a categorization of searches based on the problems that adversely affected the retrieval of potentially useful information from the experimental systems. Problems were the same as problems for the failure of searches in both systems at Earlham except for the addition of categories for conflicting user relevance assessments and navigational problems.

10.3.2.1 User Perseverance

User perseverance was again a common reason for the failure of searches to retrieve useful titles. For example, a user entered the query "roosevelt eleanor." The Blue System displayed an alphabetical list of subject headings for personal names and highlighted the closest matching name heading "Roosevelt, Eleanor, 1884–1962." The user terminated the search at this point without selecting the subject heading or displaying titles connected to it. Of the ten Blue-system searches marred by user perseverance problems, eight followed on the heels of a Pinstripe-system search for the same topic. Users could have displayed low levels of perseverance in the course of these ten searches because the results of the Blue-system searches were rather similar to the results of the Pinstripe-system searches.

Table 10.25 recounts user and system activity that resulted in attributing the failure of several searches conducted in the Blue and Pinstripe Systems to user perseverance. If users had continued searching, these searches would have most likely succeeded in retrieving useful titles.

Previously, we speculated that the results of controlled vocabulary searches could have been too complex for users to understand (sections 10.2.1.1 and 10.3.1.1). That is, they could have become lost or confused about which of the several options to pursue for continuing the search. This does not explain what was in the mind of the user who conducted the search for "united states imperialism." The Blue System retrieved a manageable number of titles (21), but the user displayed only four titles, and gave all four of them negative relevance assessments.

A total of seventeen searches failed because of user perseverance. Ten of these searches were conducted in the Blue System. The failure analysis demonstrated that these searches were really successful searches in terms of leading users to titles on their topics of interest.

Table 10.25. User Perseverance

User query	Explanation
underground railroad	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, notes, place subdivisions, and for expanding the search are enabled. The user ends the search.
black women	Pinstripe System's keyword-in-record approach retrieves 42 titles including *Black women in American history: the twentieth century, Black women in American history: from colonial times through the nineteenth century,* and *I dream a world: portraits of Black women who changed America*. The user walks away without rating any titles.
minorities	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, notes, place subdivisions, and for expanding the search are enabled. The user terminates the search.
united states imperialism	Pinstripe System's keyword-in-record search retrieves 21 titles. The user displays the first four titles and terminates the search. Titles missed include *The rising American empire, American imperialism,* and *American's road to empire: the war with Spain and overseas expansion.*

10.3.2.2 Specificity of User Queries

The problem of specificity of user queries adversely affected the retrieval of useful titles in three Blue-system searches. The users whose searches were classified in this category probably had a different (and maybe related) topic in mind but, for whatever reason, they entered queries into the experimental online catalog that were not indicative of their information needs.

Table 10.26 recounts user and system activity that resulted in attributing the failure of three searches conducted in the Blue System to specificity of user queries.

Table 10.26. Specificity of User Queries

User query	Explanation
jackson andrew jacksonian democracy	Blue System's keyword-in-record search retrieves three titles, *The rise and decline of Jacksonian democracy*, *The meaning of Jacksonian democracy*, and *Jacksonian democracy, 1829–1837*. The user gives the three titles "possibly useful" ratings.
civil war women	Blue System's keyword-in-subdivided-heading search displays the subject heading "United States — History — Civil War, 1861–1865 — Women." The user selects the heading and displays the three titles assigned this heading, *A woman's Civil War: a diary with reminiscences of the war from March 1862*, *Bonnet brigades*, and *Noble women of the North*. The user gives the three titles negative relevance assessments.
communism	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, narrower terms, notes, place and form subdivisions, and for expanding the search are enabled. The user displays ten titles and gives none positive relevance assessments.

10.3.2.3 Conflicting Relevance Assessments

The failure of three searches in the Blue System was attributed to conflicting relevance assessments. All three searches followed a search of the Pinstripe System for the same topic. When users searched the Pinstripe System, they gave one or more titles "very useful" ratings. When they searched the Blue System, they gave the same title(s) "not useful" or "possibly useful" ratings. For example, a user who conducted a Blue-system search for "black families" rated the title *The Black family in slavery and freedom, 1750–1925* "very useful" the first time she displayed this title. She switched to the Pinstripe System, conducted a Pinstripe-system search for this same query and rated the same title "not useful" the second time she displayed it. Since we did not interview users following their searches of the experimental online catalog, we did not have the opportunity to ask users why they changed their relevance assessments from a positive to a negative rating. For the purpose of the failure analysis, we counted searches in which there was a conflicting relevance assessment as successful searches

10.3.2.4 Database Failure

Three searches failed to retrieve useful titles because of database failure. The queries in these three searches were:

- jazz
- sit ins

- black church history

The former query was entered into the Blue System only. The two latter queries were entered into the Pinstripe System only. In all three cases, users retrieved titles in response to their queries but displayed titles were much more specific than the original user query. For example, the one title retrieved by the query "black church history" in a Pinstripe-system search was *Quest for equality: the life and writings of Mary Elizabeth Church Terrell, 1863–1954.* The Pinstripe System retrieved this title in a keyword-in-record search because of the occurrence of the word "church" in the title and the words "black" and "history" in the series title, viz. "Black women in United States history."

10.3.2.5 Large Retrievals

Of searches that failed to retrieve useful titles in one of the two systems users searched, only one search retrieved too many titles. This was a Pinstripe-system search for "slavery" that retrieved 390 titles. The user responded to the display of the first of 390 titles by giving the title a "not useful" rating and terminating his search of the Pinstripe System.

10.3.2.6 Vocabulary of User Queries

Searches placed in this unsuccessful search category involved user queries that were phrased in certain ways that prevented the experimental online catalogs from producing retrievals. Of searches that failed to retrieve useful titles in one of the two systems users searched, only one search was so phrased. This was a Blue-system search for "roaring twenties." The system exhausted the search trees without making a match. If the query were reduced to a single word — "twenties" — the system would retrieve titles such as *Into the twenties: the United States from Armistice to normalcy* and *Only yesterday: an informal history of the nineteen-twenties*. A one-word query such as "twenties" is not likely to retrieve many false drops in a small database limited to library materials on American history. Such a short query would be likely to retrieve many false drops in large library databases that feature encyclopedic coverage of recorded knowledge. To guide users to additional library material, systems should feature user-feedback mechanisms that prompt users for relevance assessments and use the subject headings assigned to useful titles to find such material.

10.3.2.7 Subject Searching Approach Failures

The experimental online catalog contained information on topics that interested users. However, its failure to retrieve the desired information was attributed to the particular subject searching approach selected by the system. Search approach failures were more common in the Pinstripe System than in the Blue System. The failure of fifteen searches was attributed to search approaches in the former system whereas the failure of two searches was attributed to search approaches in the latter system.

Table 10.27. Subject Searching Approach Failures in the Pinstripe System

User query	Explanation
vietnam war	The Pinstripe System fails to retrieve titles through the keyword-in-subdivided-heading search. (Blue System retrieves 8 titles in a title-keyword search and the user rates 5 titles "very useful.")
calhoun (last name), john (first name), state rights (topic)	The Pinstripe System fails to retrieve titles through the keyword-in-record search. (Blue System ignores the topic element, places the user in an alphabetical list of subject headings where he chooses "Calhoun, John C. (John Caldwell), 1782–1850," displays three of twelve retrieved titles, and gives two of three titles "very useful" ratings.)
abolitionism	The Pinstripe System fails to retrieve titles through the keyword-in-subdivided-heading search. (Blue System retrieves eight titles in a title-keyword search and rates one title "very useful.")

Table 10.27 lists three user queries and explanations for subject search approach failures in the Pinstripe System. Failures in the Pinstripe System were usually connected with the failure of the keyword-in-subdivided-heading search to retrieve subject headings bearing the words in user-entered queries.

Table 10.28. Subject Searching Approach Failures in the Blue System

User query	Explanation
women in the westward movement	The Blue System informs the user that it is unable to retrieve titles. (In a Pinstripe-system search for the same topic, the Pinstripe System responds with the alphabetical approach and places the user at the subject heading "Women pioneers." The user finds useful titles bearing this heading.)
american indian religious freedom act	The Blue System exhausts the search trees without making a match. (Pinstripe System responds with the alphabetical approach and places the user in an alphabetical list of subject headings at the cross reference "American Indians." User selects subject headings beginning with the word "Indians" and rates retrieved titles "very useful.")

Search approach failures were not as common in the Blue System. The two search approach failures that occurred in the Blue System are described in Table 10.28.

10.3.2.8 Navigational Problems

The failure of two Blue-system searches was attributed to navigational problems (Table 10.29).

Table 10.29. Navigational Problems

User query	Explanation
marxism	The Blue System informs the user that this term is a cross reference for "Communism." An exact search for "communism" results in enabled options for displaying general works, broader terms, narrower terms, notes, place and form subdivisions, and for expanding the search. The user displays titles connected with "Communism" and a narrower term "Marxian economics" and rates them "not useful." The Blue-system user seems close to finding useful titles but may be having difficulty navigating exact-approach options to find them.
pullman strike	Blue System responds to this query with an alphabetical list of subject headings bearing the cross reference "Pullman Strike, 1894." User selects the reference and the system responds by telling the user that "Chicago Strike, 1894" is the term used for this topic. System responds with the exact-search main menu with options for displaying titles and for expanding the search. User chooses latter option and the system continues searching using the original query words "pullman strike." The system fails to retrieve titles with these words because they occur in a cross reference, not in a bibliographic record.

10.3.2.9 Failure of One System Revisited

Searches that failed to retrieve useful retrievals in one system numbered 47 — about one-quarter of the searches performed in the experimental online catalog. A failure analysis of these searches demonstrated the following problems that adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failures, (5) large retrievals, (6) user query vocabulary, (7) system failures, and (8) navigational problems.

Table 10.30 reclassifies unsuccessful searches in the Blue and Pinstripe Searches. The reclassification is limited to the 47 unsuccessful searches conducted in the Blue or Pinstripe Systems that were featured in section 10.3.2.

Table 10.30. Reclassified Searches
(Single system, Earlham)

	Blue System		**Pinstripe System**	
Useful titles retrieved	**No.**	**%**	**No.**	**%**
Successful searches†	16	72.7	7	28.0
Database failures	1	4.6	2	8.0
System failures*	5	22.7	16	64.0
Total	22	100.0	25	100.0

†Barring perseverance, query specificity, conflicting relevance assessments.

*Including problems with large retrievals, subject searching approaches, user query vocabulary.

The failure analysis promoted several unsuccessful searches to successful searches. A little under three-quarters of Blue-system searches and a little over one-quarter of Pinstripe-system searches were successful in terms of guiding users to titles on the topics their queries describe. Database failures accounted for few failures. System failures were more plentiful in the Pinstripe System than in the Blue System.

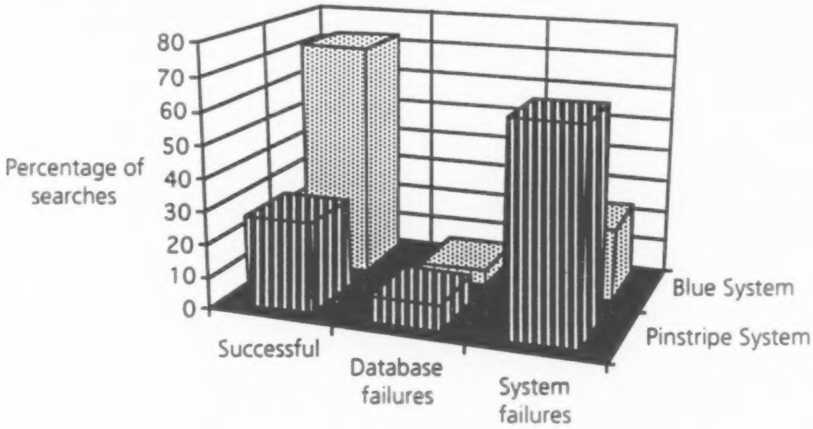


Figure 10.9. Reclassified searches
(single system, Earlham)

Figure 10.9 is a graphical depiction of the statistics in Table 10.30. It shows obvious differences in the percentages of successful searches and system failures in Blue and Pinstripe Systems. Although sample sizes were quite low in this analysis, the percentage of successful searches in the Blue System was dramatically larger than the percentage of such searches in the Pinstripe System. Furthermore, the percentage of system failures in the Pinstripe System was almost three times large as the percentage of system failures in the Blue System.

10.3.3 Revisiting Successful and Failed Searches

In the failure analysis of searches conducted by Earlham users, the researchers repeated user searches that failed to retrieve useful titles in an attempt to determine why they were unsuccessful in retrieving useful titles. In the process of repeating these searches, the failure analysis demonstrated that several searches might have been headed toward success but for various user actions that precluded the display of useful titles, e.g., perseverance, queries that were not indicative of user interests, conflicting relevance assessments. The failure analysis also demonstrated reasons for failure, e.g., design problems with a particular search approach or indexing technique, inability of a randomly-selected approach in the Pinstripe System to retrieve any titles in response to user queries. We can learn from the failure analysis in terms of designing subject searching approaches and reconfiguring the search trees. Chapter 12 makes recommendations for improving subject access in online catalogs based on failure analysis results.

Table 10.31 recasts the number and percentage of successful and unsuccessful searches based on the failure analysis of end-user searches at Earlham College.

About three-quarters of Blue-system searches were successful; about half of Pinstripe-system searches were successful. A large percentage (18.3%) of Blue-system searches would have been successful but users did not exhibit sufficient perseverance to stay with the search and see it to its conclusion. This percentage was somewhat smaller in the Pinstripe System (8.1%).

Speculating on the reasons why many Blue-system users exhibited low levels of perseverance, we are quite sure that user perseverance in the Blue System was connected with amount of time the system required for users to browse the many options available in alphabetical and exact searches. These approaches allowed users to select listed subject headings, narrower terms, broader terms, or subdivided forms of selected subject headings. It took time for users to navigate between options and for the system to search for selected subject headings. In Chapter 11, we take a close look at the alphabetical and exact approaches, identify troublesome features of these approaches, and make recommendations for improving their design in future online systems.

Table 10.31. Reclassified Successful and Unsuccessful Searches at Earlham

Type of search	Blue System		Pinstripe System	
	No.	%	No.	%
Successful searches:				
≥1 useful titles retrieved	46	44.2	37	37.8
Retrieval of useful titles marred by perseverance	19	18.3	8	8.1
Retrieval of useful titles marred by query specificity	7	6.7	3	3.1
Retrieval of useful titles marred by conflicting relevance assessments	3	2.9	0	0.0
Subtotal	75	72.1	48	49.0
Unsuccessful searches:				
Database failure	15	14.5	16	16.3
Search approach failures	2	1.9	23	23.4
Large retrievals	0	0.0	3	3.1
Vocabulary of user queries	10	9.6	8	8.2
Navigational problems	2	1.9	0	0.0
Subtotal	29	27.9	50	51.0
Total	104	100.0	98	100.0

Of the reasons for unsuccessful searches, database failure affected about the same percentages of searches conducted in the Blue and Pinstripe Systems. Vocabulary of user queries was also troublesome and led to a little under ten percent of search failures in both Blue and Pinstripe Systems. What was most striking about the percentages of search failures was the large percentage (23.4%) of failures for search approaches in the Pinstripe System. Only 1.9% of searches in the Blue System failed because of subject searching approaches. Most of the time, the Pinstripe System's keyword-in-subdivided-heading search was unable to retrieve titles bearing words in user queries. On occasion, its alphabetical approach placed users in an alphabetical neighborhood that was not fruitful for helping users find satisfactory subject headings for expressing their topics of interest.

The Pinstripe System's selection of a subject searching approach in response to a user-entered query was based on a random selection algorithm. In contrast, the Blue System's selection of a subject searching approach was governed by search trees. These data provide evidence to support the development of search trees in online catalogs. In a catalog without search trees, a substantial percentage of end-user searches are bound to fail because systems do not assist end users in subject searching approach selection.

10.4 Chapter Summary

The approach used in this study to determine why the experimental ASTUTE online catalog was successful or unsuccessful in retrieving useful titles was a failure analysis. The project director, an expert searcher of information retrieval systems, repeated end-user searches that were successful and unsuccessful in retrieving useful titles to identify the reasons why they succeeded or failed. This section features the results of the failure analysis.

The failure analysis focused on searches in which users displayed no useful titles. This was a majority of the searches conducted in the Blue and Pinstripe Systems at UM-D and Earlham (see Tables 10.1 and 10.18) as users failed to display of useful titles in almost three-fifths of their experimental online catalog searches.

The failure analysis of these searches demonstrated the following problems that adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failures, (5) large retrievals, (6) user query vocabulary, (7) search approach failures, (8) navigational problems, and (9) spelling problems. Sections 10.2 and 10.3 described the failure analysis of searches conducted by users at UM-D and Earlham, respectively. The discussion in these two sections was divided into two parts for searches that failed to produce useful titles in both experimental systems and searches that failed to produce useful titles in one of the two systems.

Explanations and examples of searches placed in categories for perseverance, query specificity, and conflicting relevance assessments provided evidence that these searches were really successful searches in terms of leading users to titles on their topics of interest. With regard to user perseverance, the systems gave users ample opportunities for browsing terminology or displaying titles on their topics of interest but users terminated their searches. In searches marred by query specificity, users gave negative relevance assessments to the titles that seemed relevant to their queries. Their actions led us to conclude that the systems were successful in terms of retrieving titles on the topics users expressed in their queries but they were not successful in terms of retrieving titles that fulfilled the specifications of queries users had in their minds but did not make explicit in their queries. In the failure analysis, we also considered searches bearing conflicting relevance assessments as successful searches because users had given the titles positive assessments in the course of their searches.

The reclassification of searches placed into categories for perseverance, query specificity, and conflicting relevance assessments from unsuccessful to successful searches resulted in success rate of about 72% for Blue-system searches at UM-D and Earlham and a success rate of between 49% and 59% for Pinstripe-system searches at UM-D and Earlham (see Tables 10.17 and 10.31). Clearly, the Blue System performed better than the Pinstripe System in terms of retrieving and displaying titles on the topics users made explicit in their queries.

Success in the Blue System, however, was marred by a large percentage of searches in which users did not persevere in terms of browsing subject terminology and displaying retrieved titles (sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1). Almost ten percentage points separated the results of Blue- and Pinstripe-system searches for searches marred by perseverance (see Tables 10.17 and 10.31). Many Blue-system searches affected by perseverance featured controlled vocabulary search approaches. Navigational problems were also linked to user perseverance and controlled vocabulary searches (sections 10.2.1.8, 10.2.2.8, and 10.3.2.8). Of the several suggestions made to improve controlled vocabulary searches, the important ones were:

- Reducing the response time for searching selected terms and creating exact-search option menus.
- Reserving exact-search option menus for summaries of large retrievals and seeking other methods for summaries of one or a few retrievals.

Success in the Pinstripe System was negatively affected by the system's random selection algorithm for subject searching approaches. Between ten and twenty percentage points separated the results of Blue- and Pinstripe-system searches for search approach failures (see Tables 10.17 and 10.31). The keyword-in-subdivided-heading search was usually to blame for search approach failures; however, the alphabetical approach had a share of the failures. Subject searching approach failures were rare in the Blue System, accounting for only 1.3% of Blue-system searches. Search approach failures accounted for the failure of 14.3% of Pinstripe-system searches. The Blue System was guided in its selection of a subject searching approach by the search trees. Thus, the data and analyses in this project provided evidence that search trees improved the search performance of subject searchers at the online catalog in terms of guiding them to subject headings and titles that described the topics they made explicit in their queries.

References

Lancaster, F. Wilfred, and Amy J. Warner. 1993. *Information retrieval today.* Arlington, Va.: Information Resources Press.

Lynch, Clifford. 1990. "Large database and multiple-database problems in online catalogs. In *Annual Review of OCLC Research, July 1989–1990,* 51–5. Dublin, Ohio: OCLC.

Markey, Karen, and Anh N. Demeyer. 1986. *Dewey Decimal Classification Online Project.* Dublin, Ohio: OCLC. OCLC Research Report Series OCLC/OPR/RR-86/1.

Wiberley, Stephen E. and Robert Allen Daugherty. 1988. "Users' persistence in scanning lists of references." *College & Research Libraries* 49, 2 (March): 149–56.

Wiberley, Stephen E., Robert Allen Daugherty, and James A. Danowski. 1993. *User persistence in displaying LUIS postings: A report to the Council on Library Resources.* Chicago: University of Illinois at Chicago. 23 July.

11 Failure Analysis of Subject Searching Capabilities

11.1 Introduction

Chapter 11 features a failure analysis of subject searching capabilities in the Blue and Pinstripe Systems. It focuses primarily on both the success and failure of these capabilities.

Table 11.1 lists Blue-system subject searches and the number of matches these capabilities made for user queries. Table 11.1 does not include matches that the Blue System made as a result of a user action to expand searches. It is divided into two portions: (1) Blue-system searches that made matches, and (2) reasons for the failure of the Blue System to effect matches.

Table 11.1. Matches in Blue-system Searches

Search approach	UM-D Blue	Earlham Blue
Searches making matches		
Exact	149	23
Alphabetical	64	12
Keyword-in-main-heading	6	0
Keyword-in-subdivided-heading	2	7
Title-keyword	52	15
Keyword in subject heading fields	2	2
Keyword-in-record (subjects)	6	5
Keyword-in-record (personal names)	0	4
Alphabetical (personal names)	0	19
Subtotal	281 (79.6%)	87 (83.7%)

No matches		
Vocabulary problems	16	10
Database failure	49	5
Search approach failures	4	2
Spelling problems	3	0
Subtotal	72 (20.4%)	17 (16.3%)
Total	353	104

Subject searching capabilities made matches for 79.6% and 83.7% of queries entered by Blue-system users at UM-D and Earlham, respectively. The exact, alphabetical, and title-keyword searches frequently made matches of user queries. Also, the Blue system's alphabetical approach frequently made matches of the personal-name queries that Earlham users entered.

The Blue System's subject searching capabilities failed to make matches for 20.4% and 16.3% of queries at UM-D and Earlham, respectively. Under the "no matches" portion of Table 11.1 are listed four reasons for search failure (i.e., vocabulary problems, database failure, search approach failure, and spelling problems). The discussion in Chapter 10 gave these same reasons for the failure of many more searches than are listed in Table 11.1. Chapter 10 failures were attributed to the failure of searches based on user relevance assessments (i.e., failed searches retrieved no titles that users rated "very useful"). Chapter 11 failures were attributed to the system's inability to produce retrievals in response to user queries.

Table 11.2 lists subject searching approaches in the Pinstripe System and the number of queries submitted to these approaches. It submitted a lot of queries for subjects generally to the keyword-in-record search.

Table 11.2. Queries Submitted to the Pinstripe System

Search approach	UM-D Pinstripe	Earlham Pinstripe
Searches making matches		
Alphabetical	53	11
Keyword-in-subdivided-heading	124	23
Keyword-in-record (subjects)	179	43
Keyword-in-record (personal names)	n/a	21
Total	356 (100.0%)	98 (100.0%)

11.2 Controlled Vocabulary Searching Based on the Exact Approach

11.2.1 Review of the Exact Approach

In the Blue System, controlled vocabulary searching focused on the exact approach. If the user-entered term matched an exact or normalized form of a controlled vocabulary term, the Blue System responded with the exact-approach main menu. The menu gave users several options for continuing their search: displaying titles assigned the matched term, displaying subdivisions connected with the matched term, browsing notes, browsing broader or narrower terms, or expanding their results with the results of a different subject searching approach.

The exact approach formed a large part of the Blue System's alphabetical approach and keyword-in-main-heading search. If the user-entered term matched a longer subject heading or *see* reference, the system responded with the alphabetical approach. This approach featured an alphabetical list of headings and references in the alphabetical neighborhood of the matched term. The system responded to users who chose a listed *see* reference with a message telling users that the system was retrieving titles assigned the particular subject heading connected with the *see* reference. It then constructed an exact-approach main menu for the particular subject heading. The system responded to users who chose a listed subject heading with the exact-approach main menu. Options in the exact-approach main menus were enabled or disabled depending on their availability for the selected subject heading.

If words in the user-entered term matched words in the main-heading subfield of subject headings, the Blue System listed matched main headings in alphabetical order. The system responded to users who chose a listed main heading with the exact-approach main menu. Options in the exact-approach main menus were enabled or disabled depending on their availability for the selected main heading.

The alphabetical and keyword-in-main-heading approaches were little more than selection mechanisms for users whose entered terms did not match subject headings or *see* references in their entirety. When users made a selection, exact-approach options enabled them to explore terminology and display titles.

The alphabetical approach was also implemented in the Pinstripe System. A random selection algorithm governed the Pinstripe System's selection of the alphabetical approach. The Pinstripe System responded with the alphabetical approach regardless of the particular elements in user queries.

11.2.2 Precision in Controlled Vocabulary Searches

Figure 11.1 shows precision scores for the three different types of controlled vocabulary searches that were designed around the exact approach. (Chapter 11 figures showing precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles because these figures isolate retrievals from particular subject searching approaches.)

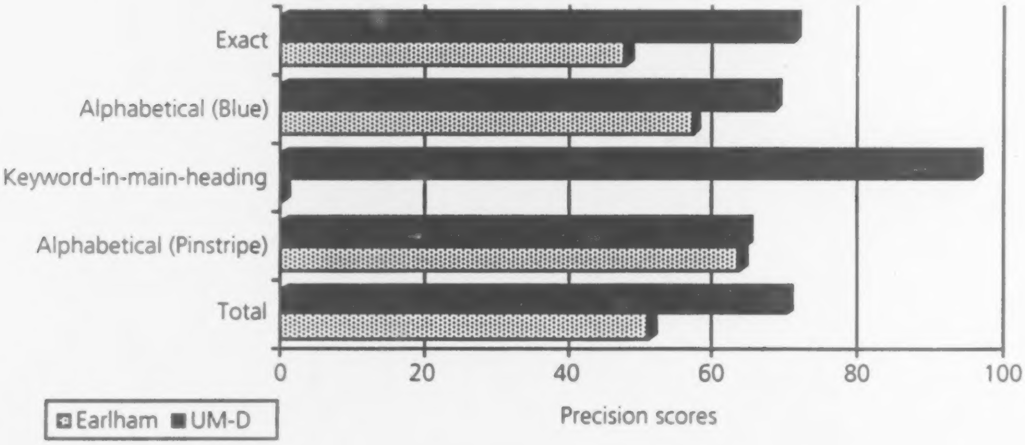


Figure 11.1. Precision in controlled vocabulary searches

Precision in controlled vocabulary searches varied considerably. They ranged from a low of 47.8% for exact searches in Earlham's Blue System to a high of 95.8% for keyword-in-main-heading searches in UM-D's Blue System. The precision score for the keyword-in-main-heading searches was quite high but it was obtained from the results of only six searches. Earlham's Blue System submitted no queries to the keyword-in-main-heading search. Generally, precision scores were higher at UM-D (70.0%) than at Earlham (51.1%).

11.2.3 Browsing and Display Opportunities in Controlled Vocabulary Searches

11.2.3.1 Browsing Options

The exact approach gave users several opportunities to browse terminology connected with the subject headings their queries matched or the subject headings they selected from alphabetical lists. Tables 11.2 and 11.3 list browsing and display opportunities in the Blue and Pinstripe Systems and the number of times users selected these options in controlled vocabulary searches at UM-D and Earlham, respectively.

UM-D users conducted 149 exact searches in the Blue System. Most browsing/display opportunities of which users took advantage were displaying titles or explicitly quitting their search. If events for displaying the exact-approach main menu, displaying titles, or quitting searches were not considered in the total number of exact-search opportunities, then searchers hardly took more than one opportunity per search. Despite the availability of broader and narrower terms and subject subdivisions for many subject headings, few users chose these options to continue their searches.

UM-D users conducted 64 alphabetical searches in the Blue System. The browsing/display opportunities they were most likely to choose were selecting terms from or browsing backward or forward in the alphabetical list, expanding their search, authorizing the system to search for a subject heading following the user's entry of a *see* reference, or explicitly quitting the search. Users displayed few broader terms, narrower terms, or subject subdivisions.

Users performed no browsing in connection with subject headings retrieved through the keyword-in-main-heading search.

Table 11.3. Browsing/Display Opportunities (UM-D)

	Blue System			**Pinstripe System**
Browsing/display opportunities	**Exact**	**Alphabetical**	**Keyword-in-main**	**Alphabetical**
Display main menu (exact approach only)	149	n/a	n/a	n/a
Display alphabetical list of subject headings (alphabetical and keyword-in-main searches only)	n/a	64	6	53
Select term or browse forward or backward in alphabetical list (alphabetical and keyword-in-main searches only)	n/a	46	4	60
Choose authorized subject headings in response to entry of *see* reference	41	24	0	25
Display titles	139	27	5	76
Browse broader terms	16	3	0	25
Browse narrower terms	32	4	0	26
Display scope note	3	0	0	5
Display topical subdivisions	24	4	0	13
Display geographic subdivisions	4	0	0	6
Display form subdivisions	14	3	0	11
Display period subdivisions	0	0	0	0
Expand search (Blue System only)	26	31	3	n/a
Redisplay original main menu or alphabetical list	8	4	0	15
Quit search	106	29	2	39
Total	562	239	20	354
Av. no. of opportunities taken per search	3.8	3.7	3.3	6.7

Generally, users conducting controlled vocabulary searches in the Blue System took fewer than four opportunities per search. They took almost seven opportunities per alphabetical search in the Pinstripe System. The browsing/display opportunities in the Pinstripe System they were most likely to choose were selecting terms from or browsing backward or forward in the alphabetical list, expanding their search, displaying titles, or explicitly quitting their search. They were also likely to display broader or narrower terms.

Table 11.4. Browsing/Display Opportunities (Earlham)

	Blue System		Pinstripe System
Browsing/display opportunities	**Exact**	**Alphabetical**	**Alphabetical**
Display main menu (exact approach only)	23	n/a	n/a
Display alphabetical list of subject headings (alphabetical approach only)	n/a	12	11
Select term or browse forward or backward in alphabetical list (alphabetical approach only)	n/a	6	17
Choose authorized subject headings in response to entry of *see* reference	7	1	1
Display titles	26	2	11
Browse broader terms	3	0	1
Browse narrower terms	6	1	1
Display scope note	2	0	1
Display topical subdivisions	6	0	5
Display geographic subdivisions	2	0	6
Display form subdivisions	2	0	1
Display period subdivisions	0	0	2
Expand search (Blue System only)	9	7	n/a
Redisplay original main menu or alphabetical list	0	0	4
Quit search	7	6	10
Total	93	35	71
Av. no. of opportunities taken per search	4.0	2.9	6.5

Earlham users conducted almost two dozen (23) exact searches in the Blue System. Most browsing/display opportunities of which users took advantage were displaying the exact-approach main menu, displaying titles, choosing authorized subject headings in response to their entry of *see* references, and expanding their searches with the results of a different subject search approach. Despite the availability of broader and narrower terms and subject subdivisions, few users chose these options to continue their searches.

Earlham users conducted a dozen alphabetical searches in the Blue System. The browsing/display opportunities they were most likely to choose were selecting from or browsing backward or forward in the alphabetical list, expanding their search, or explicitly quitting the search. Users selected few listed terms from alphabetical lists. They also hardly displayed titles, broader or narrower terms, or subject subdivisions.

Earlham users conducted only eleven alphabetical searches in the Pinstripe System. The browsing/display opportunities they were most likely to choose were displaying titles, explicitly quitting the search, and, especially, browsing backward or forward in the alphabetical list. Alphabetical searches in the Pinstripe System averaged between six and seven opportunities per search; Blue-system searches averaged four or fewer opportunities per search.

The amount of browsing UM-D and Earlham users performed in the Pinstripe System was expected because this system's selection of the alphabetical approach was not based on the extent to which listed terms matched user queries; consequently, users performed a lot of alphabetical browsing to find subject headings that described their topics of interest.

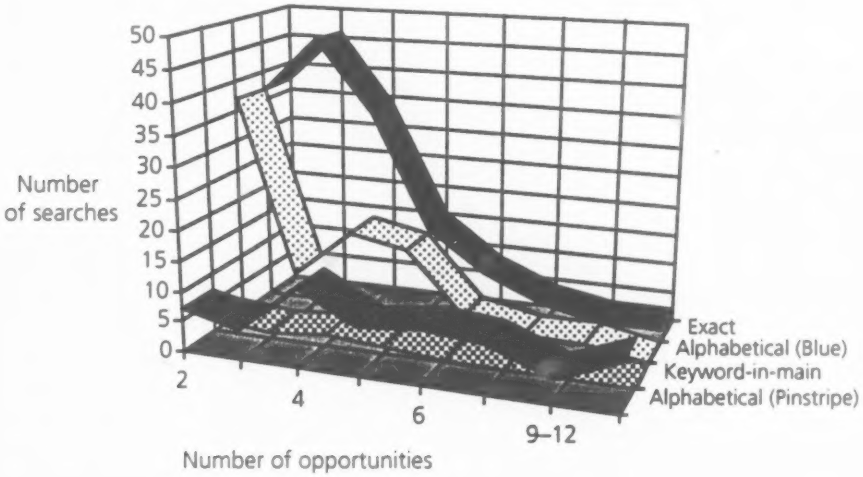


Figure 11.2. Number of opportunities per search (UM-D and Earlham)

Figure 11.2 shows the number of opportunities users took per search. Although users took few opportunities in controlled vocabulary searches, they took even fewer opportunities in Blue-system searches than in Pinstripe-system searches.

Generally, users conducting exact searches took two to four opportunities per search. Only four exact searches exceeded eight opportunities. The longest exact search featured 24 opportunities. This search was initiated by the query "fluids." The user browsed over a half dozen narrower terms, two broader terms, displayed titles assigned these broader and narrower terms, and displayed topical subdivisions connected with a few of these terms. Generally, users conducting alphabetical searches in the Blue System took two to five opportunities per search. Keyword-in-main searches were also quite short; they ranged from two to five opportunities per search.

Several alphabetical searches in the Pinstripe System were quite long; eight searches featured thirteen or more browsing/display opportunities. Although the Pinstripe System randomly chose the alphabetical search in these eight searches, all eight searches benefited from terms in the alphabetical list that described their topics of interest. The query "composites" initiated the longest search that featured 34 browsing/display opportunities. The user selected the term "Composite materials" from the alphabetical list. The user browsed eight narrower terms, a half dozen broader terms, a few subject subdivisions, and displayed titles assigned to some terms and subdivided subject headings.

Lengthy searches that featured many browsing/display opportunities were the exception rather than the rule in the experimental online catalog's controlled vocabulary searches. Surprised by the small amount of browsing users did in the Blue and Pinstripe Systems, the researchers performed a failure analysis of successful controlled vocabulary searches to shed light on browsing in the experimental online catalogs. They divided successful controlled vocabulary searches into the following four categories:

- Simple, successful searches — searches in which users performed no browsing to find useful titles. Users only selected the <display general works> option from the exact-approach main menu, and displayed and rated one or more titles.
- Successful searches with a little browsing — searches in which users selected one browsing option from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term.
- Successful searches with more than a little browsing — searches in which users selected two or more browsing options from the exact-approach main menu in addition to displaying titles connected with the browsed term or the original matched term.

- Successful searches due to expanding — searches users conducted in the Blue System in which they expanded the search to find useful titles.

Table 11.5. Browsing in Successful Controlled Vocabulary Searches

Search category	**UM-D**		**Earlham**	
	Blue	**Pinstripe**	**Blue**	**Pinstripe**
Simple search				
Exact	55	n/a	7	n/a
Alphabetical	14	10	0	2
Keyword-in-main	3	n/a	0	n/a
Subtotal	72 (59.5%)	10 (37.1%)	7 (50.0%)	2 (40.0%)
A little browsing				
Exact	12	n/a	4	n/a
Alphabetical	1	9	0	3
Keyword-in-main	0	n/a	0	n/a
Subtotal	13 (10.8%)	9 (33.3%)	4 (28.6%)	3 (60.0%)
More than a little browsing				
Exact	12	n/a	1	n/a
Alphabetical	2	8	0	0
Keyword-in-main	1	n/a	0	n/a
Subtotal	15 (12.4%)	8 (29.6%)	1 (7.1%)	0 (0.0%)
Expanded search				
Exact	5	n/a	1	n/a
Alphabetical	15	n/a	1	n/a
Keyword-in-main	1	n/a	0	n/a
Subtotal	21 (17.3%)	n/a	2 (14.3%)	n/a
Total	121 (100%)	27 (100%)	14 (100%)	5 (100%)

These four successful-search categories are listed in Table 11.5. Each category is further subdivided by the particular subject searching approach that the Blue or Pinstripe System invoked to find a match.

Overall, simple searches made up over half of successful searches. Expanded searches were somewhat common in the Blue System where they accounted for about one of every eight searches.

Sections 11.2.3.2–11.2.3.6 describe the failure analysis of browsing in successful controlled vocabulary searches.

11.2.3.2 Simple, Successful Controlled Vocabulary Searches

In simple, successful controlled vocabulary searches, users performed almost no browsing to find useful titles. Of the 167 successful controlled vocabulary searches, 91 (54.5%) were simple, successful controlled vocabulary searches.

Table 11.6 describes simple, successful controlled vocabulary searches in the Blue and Pinstripe Systems.

Table 11.6. Simple, Successful Controlled Vocabulary Searches

User query	Explanation
cad	Blue System informs the user that the entered term is a cross reference to "Computer-aided design." It responds with the results of an exact search in which options for displaying general works, narrower terms, subject and form subdivisions, and for expanding the search are enabled. User displays 2 of 22 retrieved titles and rates both "very useful."
cast iron	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, subject subdivisions, and for expanding the search are enabled. User displays all 5 retrieved titles and rates them "very useful."
american indian religious freedom act	Pinstripe System responds with the alphabetical approach in which "American Indians" is the closest matching term. User selects this *see* reference for "Indians of North America," retrieves 12 titles, and rates one title useful.
c language	Blue System responds with results of a keyword-in-main-heading search in which three subject headings contain the same words as the user query (i.e., "C (Computer program language)," "PL/C (Computer program language)," and "Small-C (Computer program language)." User selects the former, displays all 42 retrieved titles, and rates all but one title "very useful."

11.2.3.3 Successful Controlled Vocabulary Searches with a Little Browsing

Successful controlled vocabulary searches that were characterized by "a little browsing" were searches in which users selected one browsing option from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term or expanding the search.

Table 11.7 describes successful controlled vocabulary searches with a little browsing.

Table 11.7. Successful Controlled Vocabulary Searches With a Little Browsing

User query	Explanation
atomic bomb	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, topical subdivisions, and for expanding the search are enabled. User browses topical subdivisions, selects "History," displays one title and rates it "very useful."
operating systems	Blue System responds with the results of an exact search in which options for displaying general works, narrower terms, topical and form subdivisions, and for expanding the search are enabled. User retrieves 32 titles for "Operating systems," displays 17 titles, and rates 5 titles "very useful." User displays narrower terms, chooses "Distributed operating systems (Computers)," and rates the 1 retrieved title (*Open systems)* "very useful."
women in the westward movement	Pinstripe System responds with the alphabetical approach in which "Women pioneers" is the closest matching term. User selects this heading, chooses geographical subdivisions connected with the heading, and retrieves three titles which she rates "very useful."
heat exchangers design	Pinstripe System responds with the alphabetical approach in which "Heat insulating materials" is the closest matching term. User chooses preceding term "Heat exchangers." Exact options for displaying titles, browsing broader or narrower terms, topical and form subdivisions, and expanding the search are enabled. User displays the 13 retrieved titles and gives them negative ratings. He chooses topical subdivisions option, selects "Design and construction," and rates the two retrieved titles (*Industrial heat exchangers* and *Heat exchangers: selection, design, and construction*) "very useful."

Of the four successful searches described in Table 11.7, two searches were conducted in the Pinstripe System. Although the terminology of user queries and displayed subject headings differed, users found suitable expressions for their topics of interest by browsing the alphabetical list of subject headings and/or by browsing subdivided forms of selected subject headings.

11.2.3.4 Successful Controlled Vocabulary Searches with More than a Little Browsing

Successful controlled vocabulary searches that were characterized by "more than a little browsing" were searches in which users selected more than one browsing option from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term or expanding the search.

Three successful controlled vocabulary searches with more than a little browsing are described in Table 11.8.

Table 11.8. Successful Controlled Vocabulary Searches With More than a Little Browsing

User query	Explanation
medical	Pinstripe System responds with the results of an alphabetical search in which "Medical care — Moral and ethical aspects" is the closest matching term. During this lengthy search, the user chooses several terms from the alphabetical list, e.g., "Medical statistics," "Medical writing," "Medicine," displays broader terms, narrower terms, notes, subject subdivisions, and titles.
heat transfer	Blue system informs the user that the entered term is a cross reference to "Heat — Transmission." It responds with the results of an exact search. User displays 42 titles assigned to the subject heading and gives all titles negative ratings. User displays narrower terms, selects "Heat exchangers" from a list of five terms, displays 13 titles, rates several useful. User also finds useful records under this subject heading that are subdivided by "Design and construction."
camera	Blue System responds with the results of an alphabetical search in which "Camera lenses," "Camera operators, Motion pictures," and "Cameras" begin with the same word as the user query. User chooses the former which is a *see* reference to "Lenses, Photographic," rates all 7 retrieved titles "very useful." User selects more terms from alphabetical list, e.g., "Photography," "Cameras," and rates titles useful.

11.2.3.5 Successful Controlled Vocabulary Searches with Expanding

Successful controlled vocabulary searches that were characterized by expanding were searches users conducted in the Blue System in which they expanded the search to find useful titles.

Table 11.9 describes four successful controlled vocabulary searches with expanding.

Table 11.9. Successful Controlled Vocabulary Searches With Expanding

User query	Explanation
psychology	Blue System responds with the results of an alphabetical search in which "Psychology" is the closest matching term. User selects this heading and the system displays enabled exact-search options for displaying narrower terms, topical subdivisions, and for expanding the search. User chooses the expand option and the system retrieves 18 titles from a title-keyword search. User displays 18 titles and gives 4 titles "very useful" ratings.
mechanical engineering	Blue system informs the user that the entered term is a cross reference to "Mechanics, Applied." In response to exact-search main menu options, user selects expand option and system responds with the results of an alphabetical search. User chooses "Mechanics, Dynamic" from the alphabetical list, retrieves one title, and gives it a "very useful" rating.
flexible manufacturing systems	Blue System responds with the results of an exact search. In response to exact-search main menu options, user selects expand option and system responds with the results of an alphabetical search. User chooses "Flexible manufacturing systems" from the alphabetical list, retrieves 14 titles, and gives them a "very useful" rating.

Since many experimental online catalog searchers were using the system for the first time, some searchers might have thought that selecting the Blue System's "expand" option was the way to display titles connected with a selected subject heading. For example, in the "flexible manufacturing systems" search, the searcher disregarded the "display general works" option, selected the "expand this search" option, and completed the alphabetical search with search options that would have been the same for completing the exact search for the same subject.

Future implementations of an expand capability should feature experimentation with users to ensure that understandable terminology is used to indicate options for displaying titles and expanding the search with the results of a different subject searching approach.

11.2.3.6 Discussion of Browsing in Successful Controlled Vocabulary Searches

Despite the browsing opportunities that controlled vocabulary searches offered, end users performed little browsing. Users seemed satisfied with displaying titles under the "general works" option and terminating the search. More browsing characterized alphabetical searches, especially in the Pinstripe System. The Pinstripe System responded randomly with the alphabetical approach regardless of the extent of match between user queries and subject headings. (The Blue System responded with the alphabetical approach when user queries matched the initial words of subject headings in the alphabetical list.) Perhaps users sensed that their terms were not exact matches of subject headings and were predisposed to browse using enabled opportunities. Interestingly, the lengthy browsing that characterized a handful of Pinstripe-system searches was initiated by queries that were exact or almost exact matches of controlled vocabulary terms.

When users initiated searches that matched controlled vocabulary terms, they were unlikely to browse for additional terminology or refine their original terminology with more specific terminology that narrower terms and subject subdivisions could have provided. Perhaps the searchers that really need browsing opportunities are those whose terms are not close to or controlled vocabulary terms. In the absence of post-search interview data that explores browsing in subject searches, we can only speculate on the paucity of browsing in controlled vocabulary searches. This may also be a fruitful area for additional research.

11.2.4 Perseverance in Controlled Vocabulary Searches

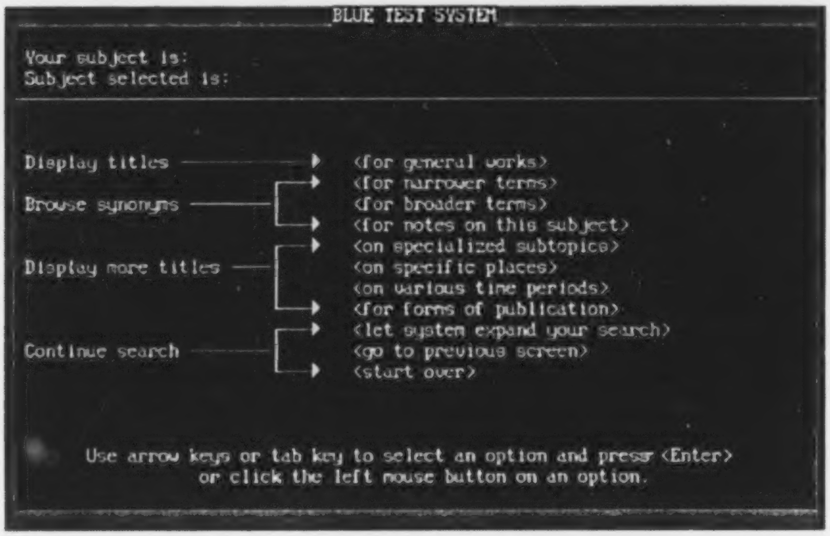
The failure analysis in chapter 10 featured discussions of perseverance in subject searches (sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1). Perseverance was a major problem in controlled vocabulary searches. Of the 272 controlled vocabulary searches in the Blue and Pinstripe Systems at UM-D, retrieval in 63 (23.2%) searches was marred by user perseverance. Of the 46 controlled vocabulary searches in the Blue and Pinstripe Systems at Earlham, retrieval in eight (15.4%) searches was marred by user perseverance.

Table 11.10 describes four searches that were adversely affected by user perseverance. These four searches were typical examples of controlled vocabulary searches in the Blue and Pinstripe Systems in which the systems were successful in matching user queries with subject headings that described users' topics of interest. Unfortunately, users did not display titles assigned those subject headings.

Table 11.10. Controlled Vocabulary Searches Marred by User Perseverance

User query	Explanation
turbo prolog	Blue System responds with the results of an exact search in which exact-approach main menu options for displaying general works, broader terms, and expanding the search are enabled. User quits. Had she selected the "general works" option, she would have retrieved one title, *Introduction to Turbo prolog.*
underground railroad	Blue System responds with the results of an exact search in which exact-approach main menu options for displaying general works, broader terms, notes, geographic subdivisions, and expanding the search are enabled. User quits. Had he selected the "general works" option, he would have retrieved several titles on this topic, e.g., *The Underground Railroad: from slavery to freedom, The liberty line: the legend of the Underground Railroad.*
automatic transmission	Blue system responds with the results of an alphabetical search in which "Automatic transmissions, Automobile" is the closest match. The user selects it and the system informs the user that the selected term is a cross reference to "Automobiles — Transmission devices, Automatic." The system responds with the results of an exact search with options for displaying general works and expanding the search. User switches systems and misses potentially useful titles such as *Automatic transmission* and *Automatic and fluid transmissions.*
c programming language	Pinstripe system responds with the results of an alphabetical search in which "C (Computer program language)" is the closest match. The user selects it and the system responds with the results of an exact search with options for displaying general works, broader and narrower terms, and expanding the search. User switches systems and misses several potentially useful titles.

In the absence of post-search interview data that probed for reasons why users terminated these seemingly successful searches, we can speculate on such reasons. One reason may be the density or complexity of the exact-approach main menu on which was listed options for displaying titles and browsing subject terms. Figure 11.3 shows the exact-approach main menu that the Blue System displayed to users in exact searches and exact-search components of alphabetical and keyword-in-main-heading searches.



```
                              BLUE TEST SYSTEM
Your subject is:
Subject selected is:

Display titles ─────────────▶  <for general works>
                         ┌──▶  <for narrower terms>
Browse synonyms ─────────┤     <for broader terms>
                         └──▶  <for notes on this subject>
                         ┌──▶  <on specialized subtopics>
Display more titles ─────┤     <on specific places>
                         │     <on various time periods>
                         └──▶  <for forms of publication>
                         ┌──▶  <let system expand your search>
Continue search ─────────┤     <go to previous screen>
                         └──▶  <start over>

      Use arrow keys or tab key to select an option and press <Enter>
               or click the left mouse button on an option.
```

Figure 11.3. Exact-approach main menu

The researchers were limited in the design of the exact-approach main menu by the FoxPro database management system. After considerable discussion amongst project team members and consultants, we accepted the design of the menu in figure 11.3 in which phrases on the left side of the menu described possible next steps in the search. These phrases pointed to specific options on the right side of the menu. Options were enabled or disabled depending on the particular matched subject heading. They were also "hot" in that selection of a particular option provoked the system to perform the described action.

The menu was dense and complex. The screen contained mostly text with little in the way of graphics except for arrows leading from phrases to specific options. For some users, this menu could have been so dense or complex that they chose to start over rather than study or experiment with listed options. Other users might not have understood that the "display general works" option would have resulted in a display of retrieved titles for the matched subject heading. Section 11.2.3.5 described several searches in which the logical next step would have been for users to choose the "display general works" option, but, instead, users chose the "expand search" option. Perhaps they did not understand the meaning of the "display general works" option and/or "expand search" option.

The design of the exact-approach main menu is critical to the exact search approach's success in retrieving titles on users' topics of interest. Future implementations of the

BEST COPY AVAILABLE

exact search must include end users in the design and pretesting of this menu, especially in the wording of listed options.

11.2.5 Navigation Problems in Controlled Vocabulary Searches

The failure analysis in chapter 11 featured a general discussion of navigation in subject searches (sections 10.2.1.8, 10.2.2.8, and 10.3.2.8). Navigation was not an obvious problem in controlled vocabulary searches because we did not conduct post-search interviews with end users to probe reasons for search failures. The researchers inferred navigation problems based on the illogical or strange actions users took in the course of their searches. For example, in response to the user-entered query "plastic molding," the Pinstripe System responded with the alphabetical approach in which the term "Plastic molding" was the closest match of the query. When the user selected this term, the system informed him that it was a *see* reference for "Plastics — Molding" and responded with an exact-approach main menu with enabled options for displaying titles, broader and narrower terms, form subdivisions, and expanding the search. The user browsed narrower terms, selected "Injection molding of plastics" and displayed subdivided forms of this subject heading. Instead of displaying the original alphabetical list bearing the term that described his query, i.e., "Plastic molding" and displaying titles on this topic, the user terminated the search. Most likely, the user did not understand how to backtrack in the search and terminated the search with the intention of starting over.

Of the 272 controlled vocabulary searches in the Blue and Pinstripe Systems at UM-Dearborn, 15 (5.5%) searches exhibited likely navigation problems. Of the 46 controlled vocabulary searches in the Blue and Pinstripe Systems at Earlham, six (13.0%) searches exhibited likely navigation problems. Although these are small percentages, many of the searches marred by perseverance might have actually been marred by navigation problems.

Table 11.11 describes four searches in which navigation problems contributed to the users' failure in retrieving useful titles.

Table 11.11. Navigation Problems in Controlled Vocabulary Searches

User query	Explanation
common business oriented language	Blue system informs the user that the entered term is a cross reference for "COBOL (Computer program language)." In response to exact-search main menu options, user selects broader-term option and selects the listed broader term "English language." User browses form subdivisions and chooses "Dictionaries." User browses topical subdivisions, e.g., "Arabic," "French," "German," and quits searching. The user's selection of the broader term "English language" dealt the killing blow to this search because it lead him away from his topic of interest.
animation	Blue system responds with the results of an alphabetical search in which "Animation, Computer" is the closest match. User browses forward in the alphabetical list, selects "Casting, Steel" which is a *see* reference for "Steel founding." User quits the search.
african-americans	Blue system informs the user that the entered term is a cross reference for "Afro-Americans." In response to exact-search main menu options, user selects narrower-term option and selects the narrower term "United States — Civil War, 1861–1865 — Afro-American influences." User quits the search. Most likely, user wants to backtrack to the exact-search main menu to explore other terms or display titles but does not know what options to pick to return to this menu.
civil rights	Pinstripe system responds with the results of an alphabetical search in which "Civil rights" is the closest match. In response to exact-search main menu options, user browses geographic subdivisions, chooses "Florida." In response to exact-search main menu options, user browses geographic subdivisions, chooses "Saint Augustine." In response to exact-search main menu options, user browses topical subdivisions, chooses "History." In response to exact-search main menu options, user browses geographic subdivisions, chooses "20th century." In the post-search interview that followed this librarian's search, he comments on the inefficiency of intervening exact-search main menus.

The navigation problems described in this section call for the following improvements to the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.

- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

11.2.6 Alphabetical Approach as a Last Resort

Although the Pinstripe System's selection of the alphabetical search was done by a random selection algorithm, there were several searches in which the closest matching term or nearby terms in the alphabetical list were not close matches in terms of syntax but users eventually found useful titles by selecting them. Several examples are given in Table 11.12.

Table 11.12. Queries and Matching Terms in Pinstripe Alphabetical Searches

User queries	Close matches in meaning
american indian religious freedom act	"American Indians" (*see* reference for "Indians of North America")
women in the westward movement	"Women pioneers"
florida key	"Florida"
computers made simple	"Computers"
emissions	System placement at "Employee incentives;" user browses forward and finds "Emission control devices" (*see* reference for "Pollution control devices").
heat exchangers design	System placement at "Heat insulating materials;" user browses forward and backward, finds "Heat exchangers," and, eventually, finds subdivision "Design and construction."

In view of users' experiences with the Pinstripe System's alphabetical approach, the researchers recommend this approach as a last resort. That is, when user queries fail to meet the criteria for controlled vocabulary or free-text approaches, search trees should force systems to respond with the alphabetical approach regardless of the extent of match between user queries and subject terms in the alphabetical list. On occasion, listed subject terms may describe the topics users have in mind and provide them with the opportunity to browse forward or backward in the list to find suitable terms.

11.3 Keyword-in-Subdivided-Heading Search

11.3.1 Review of the Keyword-in-Subdivided-Heading Search

Both Blue and Pinstripe Systems featured the keyword-in-subdivided-heading search. Like the exact, alphabetical, and keyword-in-main-heading search, this search was a controlled vocabulary search. Unlike the exact, alphabetical, and keyword-in-main-heading search, the user's entry of a subject query or selection of a system-displayed subject heading did not result in the exact approach. Instead, the systems responded with an alphabetical list of main and subdivided subject headings bearing the same words as those in user queries. The user's selection of a listed subject heading resulted in a display of titles assigned the selected subject heading. When users terminated the display of titles, the systems returned to the display of matching subject headings so that users could select more subject headings of interest and display more titles.

The Blue and Pinstripe Systems responded to user queries with the keyword-in-subdivided-heading search under different circumstances. The Blue System responded with keyword-in-subdivided-heading search results when the words in user-entered queries matched subject headings bearing the same words, and, usually, additional words. For example, the user query "afro americans and west" invoked the keyword-in-subdivided-heading search in the Blue System. It responded with two subdivided subject headings: "Afro-Americans — West (U.S.)" and "Afro-Americans — Middle West." Both subject headings contained the same words as the words in user queries, i.e., "afro," "americans," and "west," and additional words, i.e., "us" and "middle."

A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The Pinstripe System responded with this search regardless of the particular elements in user queries.

Table 11.13 summarizes the reasons for success or failure of keyword-in-subdivided-heading searches.

Table 11.13. Keyword-in-subdivided-heading Searches

Reasons for success/failure	UM-D		Earlham	
	Blue	Pinstripe	Blue	Pinstripe
Successful searches:				
Simple search	1	25	2	1
A little browsing	0	9	0	3
More than a little browsing	0	1	0	0
Expanded search	1	n/a	1	n/a
Subtotal	2 (100.0%)	35 (28.2%)	3 (42.9%)	4 (17.4%)
Unsuccessful searches:				
Perseverance	0	24	0	0
No matches	0	28	0	12
Database failure	0	23	2	5
Query specificity	0	5	2	0
Too many retrievals	0	4	0	0
Search approach failures	0	5	0	0
Vocabulary	0	0	0	2
Subtotal	0 (0.0%)	89 (71.8%)	4 (57.1%)	19 (82.6%)
Total	2 (100%)	124 (100%)	7 (100%)	23 (100%)

Keyword-in-subdivided-heading searches were much more frequent in the Pinstripe System than in the Blue System. We attributed the search trees as the reason for this. A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The random selection algorithm had three subject searching approaches from which to choose. Thus, the Pinstripe System should have responded to about a third of user-entered queries with the keyword-in-subdivided-heading search. Search trees governed the Blue System's selection of a particular subject searching approach. They favored exact, alphabetical, and keyword-in-main-heading searches before the keyword-in-subdivided-heading search. Although users frequently entered queries into the Blue System that matched words in subject headings, this system responded to many of these queries with the exact, alphabetical, or keyword-in-main-heading searches. Blue-system users could still obtain the results of keyword-in-subdivided-heading searches for such queries by repeatedly expanding their searches.

Overall, unsuccessful keyword-in-subdivided-heading searches outnumbered successful keyword-in-subdivided-heading searches. Almost three-quarters (112 of

156) of keyword-in-subdivided-heading searches were unsuccessful. Frequent reasons for failure were perseverance, database failure, and keyword-in-subdivided-heading searches that retrieved no subject headings in the Pinstripe System.

This section describes keyword-in-subdivided-heading searches in detail, focuses on strengths of this search, and makes suggestions for improving its weaknesses.

11.3.2 Precision in Keyword-in-subdivided-heading Searches

Figure 11.4 shows precision scores for keyword-in-subdivided-heading searches. (Precision scores included conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

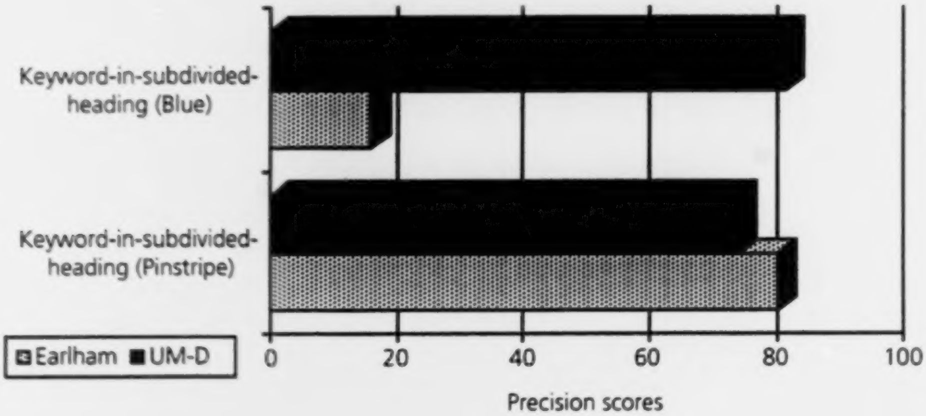


Figure 11.4. Precision in keyword-in-subdivided-heading searches

Except for precision in Earlham's Blue System, precision scores for keyword-in-subdivided-headings searches were quite high. Precision scores in Earlham's Blue System were based on the results of only seven searches.

11.3.3 Successful Keyword-in-subdivided-heading Searches

Table 11.13 lists four categories of successful keyword-in-subdivided-heading searches. Descriptions of these four categories are:

- Simple searches — searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles.
- A little browsing — searches in which users selected two listed subject headings, displayed one or more titles under each heading, and gave "very useful" ratings to one or more titles in the course of the search.
- More than a little browsing — searches in which users selected more than two listed subject headings, displayed one or more titles under more than two subject headings, and gave "very useful" ratings to one or more titles in the course of the search.
- Successful searches due to expanding — searches users conducted in the Blue System in which they expanded the search to find useful titles.

Successful keyword-in-subdivided-heading searches numbered 44. The average number of subject headings retrieved in these successful searches was 10.6. The range of retrieved subject headings was from 1 to 121. If postings for the one search that resulted in 121 subject headings were disregarded in the calculation of the average number of retrieved subject headings, this average would drop to eight subject headings per successful search. Nine keyword-in-subdivided-heading searches retrieved only one subject heading.

Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing. These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing. Table 11.14 gives examples of the four types of successful keyword-in-subdivided-heading searches.

Users performed little browsing in keyword-in-subdivided heading searches. The design of this search encouraged browsing because it always returned to the list of retrieved subject headings following the display of retrieved titles. Instead of choosing another subject heading, most users chose to terminate the search.

Table 11.14. Successful Keyword-in-subdivided-heading Searches

User query	Explanation
civil war historiography	Simple search in which the Blue System retrieves one subdivided subject heading, "United States — History — Civil War, 1861–1865 — Historiography." User selects this heading, retrieves six titles, displays all six, and rates one "very useful."
inventions	Search that features "a little browsing" in which Pinstripe System retrieves a total of seven subject headings (i.e., "Inventions," and six subdivided forms of "Inventions"). The user chooses "Inventions — Congresses" and "Inventions," retrieves 13 titles, and displays one title under each heading, and rates one title "very useful."
polymers	In the one search that features "more than a little browsing," Pinstripe System retrieves a total of 12 subject headings. The user chooses "Polymers," and five subdivided forms of this heading, e.g., "Handbooks, manuals, etc.," "Analysis," "Information services," and rates most displayed titles "very useful."
numerical control	Blue System retrieves 4 subject headings. User chooses "Machine design — Numerical control," retrieves 24 titles, and displays a few. User expands search, retrieves 32 titles in a title-keyword search, and displays some.

11.3.4 Unsuccessful Keyword-in-subdivided-heading Searches

There were many more unsuccessful keyword-in-subdivided-heading searches than successful ones. Of the seven reasons for failure, three reasons accounted for over four-fifths of failed searches: (1) searches that retrieved no subject headings in the Pinstripe System (35.7% of failures), database failure (26.8% of failures), and perseverance (21.4% of failures). This section focuses on these and other reasons for unsuccessful keyword-in-subdivided heading searches.

11.3.4.1 No Postings in Pinstripe-system Searches

The most frequent reason for failure of the keyword-in-subdivided-heading search was the failure of this search to retrieve subject headings in the Pinstripe System. Examples of user queries that failed to retrieve subject headings are:

- american events
- civil rights movement
- abolitionism

- flow charting
- software management
- computing dictionary

Although the Pinstripe System's keyword-in-subdivided-heading search failed to retrieve subject headings for these queries, subject searching approaches in the Blue System were successful retrieving subject headings and useful titles. For example, the Blue System retrieved two titles, *Dictionary of computing and new information technology* and *Dictionary of computing*, that the user rated "very useful" in a title-keyword search.

Table 11.15. Blue-system Search Outcomes for Failed Keyword-in-subdivided-heading Searches

Outcome	UM-D	Earlham
Successful search outcomes		
Title-keyword	8	6
Matching *see* reference in exact search	4	0
Keyword-in-record	2	1
Keyword in subject heading fields	0	1
Exact approach	0	1
Alphabetical approach	2	0
Unsuccessful search outcomes		
Perseverance	2	2
Query specificity	4	1
No matches (search trees exhausted)	2	0
Vocabulary problems	4	0
Total	28	12

Table 11.15 lists the subject searching approaches that the Blue System enlisted for keyword-in-subdivided-heading searches that failed in the Pinstripe System. It also gives the outcome of Blue-system searches.

A lot of keyword-in-subdivided-heading searches that failed in the Pinstripe System were successful in the Blue System through title-keyword searches. A few failed searches were successful in the Blue System because the user query matched a *see* reference in an exact search. Examples were the user queries "database management" and "ergonomics" that were *see* references for "Data base management" and "Human engineering," respectively. The availability of *see* references for these queries made it possible for users to retrieve titles on their topics of interest.

Other keyword searches figured into the success of a few searches in the Blue System, i.e., keyword-in-record and keyword in subject heading fields searches. Two user queries, "medical statistic" and "power tool" matched the initial characters in longer subject headings, "Medical statistics" and "Power tools," respectively. These two queries triggered the Blue System's alphabetical approach and, eventually, resulted in the retrieval of useful titles.

Generally, the success of searches that failed to produce retrievals through the Pinstripe System's keyword-in-subdivided-heading approach usually succeeded through the Blue System's free-text approaches, i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches, or *see* references featured in controlled vocabulary searches.

11.3.4.2 Database Failure

The failure of the keyword-in-subdivided-heading approach was frequently attributed to database failure. Examples of queries that failed because of database failure were:

- portable data file
- radar detectors
- computer workstation
- women in reconstruction
- dayton history
- minorities in the west

Database failure was to blame for the failure of users' corresponding searches of the Blue System for the same topic for many of these queries.

11.3.4.3 Perseverance

Perseverance was a problem with keyword-in-subdivided-heading searches, especially in UM-D's Pinstripe System where the failure of two dozen searches was attributed to perseverance. Perseverance could not be blamed on high numbers of retrievals because retrievals were not much different from retrievals for successful searches. For keyword-in-subdivided-heading searches marred by perseverance the average number of subject headings retrieved per query was 13.67. The range of retrieved subject headings was from 1 to 144. If postings for the one search that retrieved 144 subject headings were disregarded in the calculation of the average number of displayed subject headings, this average would drop to eight subject headings per search. This average was the same as the average for successful controlled vocabulary searches minus the one search with the highest postings. Two keyword-in-subdivided-heading searches retrieved one subject heading.

Table 11.16 describes keyword-in-subdivided-heading searches that failed due to user perseverance. This table describes seemingly successful searches. Retrievals were quite manageable in that users did not have to browse more than two screens of retrieved subject headings to find one(s) that interested them. Unfortunately, instead of browsing or selecting listed subject headings, users terminated their searches. Perhaps the low level of perseverance in keyword-in-subdivided-heading searches was indicative of the overall low level of perseverance in controlled vocabulary searches (see section 11.2.4) and low levels of browsing in controlled vocabulary searches (see section 11.2.3) and in keyword-in-subdivided-heading searches (see section 11.3.3).

Table 11.16. Perseverance in Keyword-in-subdivided-heading Searches

User query	Explanation
error correcting	Pinstripe System retrieves 1 subdivided subject heading, "Error-correcting codes." User terminates the search.
robotics	Pinstripe System retrieves 12 subject headings, "Robotics" and 11 subdivided forms of this heading, e.g., "Research," "Mathematics," "Economic aspects." User terminates the search.
composite materials	Pinstripe System retrieves 9 subject headings, "Composite materials" and 8 subdivided forms of this heading, e.g., "Fatique," "Fracture," "Testing." User terminates the search.
computer science	Pinstripe System retrieves 15 subject headings, "Computer science," and several headings with this phrase as a qualifier, e.g., "Natural language processing (Computer science)," "Neural networks (Computer science)," "Expert systems (Computer science)." User terminates the search.

11.3.4.4 Large Retrievals

Of the remaining reasons for failure, large retrievals were especially interesting. The failure of only four searches was attributed to large retrievals. User queries and retrievals in these searches were:

- Mathematics (52 subject headings)
- Computers (117 subject headings)

- Computer (over 400 subject headings)
- Materials (121 subject headings)

The experimental online catalogs did not count up retrieved subject headings and report the number to users. In searches for these four queries, users might have browsed forward and/or backward in the subject headings lists. When they realized how lengthy subject heading displays were, they terminated their searches rather than select listed subject headings.

11.4 Title-keyword Searches

Title-keyword searches were unique to the Blue System. The Blue System's search trees submitted one-word queries to title-keyword searches following unsuccessful exact and alphabetical searches or in response to the user's selection of the expand option following successful exact and/or alphabetical searches. It submitted queries exceeding one word to title-keyword searches following unsuccessful exact, alphabetical, keyword-in-main-heading, and keyword-in-subdivided-heading searches or in response to the user's selection of the expand option following successful exact, alphabetical, keyword-in-main-heading, and/or keyword-in-subdivided-heading searches. The Blue System retrieved titles that contained the same words in user queries and displayed the first of one or more retrieved records to searchers.

Table 11.17 summarizes the number of title-keyword searches and reasons for success or failure.

Table 11.17. Title-keyword Searches

	UM-D	Earlham
Reasons for success/failure	**Blue**	**Blue**
Successful searches:		
Simple search	28	9
Expanded search	8	4
Subtotal	36 (69.2%)	13 (86.7%)
Unsuccessful searches:		
Perseverance	8	1
Database failure	5	0
Query specificity	3	1
Subtotal	16 (30.8%)	2 (13.3%)
Total	52 (100%)	15 (100%)

Title-keyword searches were not very common in the Blue System. This was due to the several searches preceding title-keyword searches (i.e., exact, alphabetical, keyword-in-main-heading, and keyword-in-subdivided-heading) to which the search trees submitted user queries.

Precision was not as high in title-keyword searches as in controlled vocabulary searches. Precision scores for title-keyword searches were 47.5% and 23.3% at UM-D and Earlham, respectively.

Successful title-keyword searches outnumbered unsuccessful title-keyword searches. Almost 70% of title-keyword searches in UM-D's Blue System were successful; over 85% of title-keyword searches in Earlham's Blue System were successful. Title-keyword searches were low posted. They averaged 6.3 titles and ranged from zero to 68 titles.

On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. Users who want additional material on their topics of interest could benefit from the system's retrieval of titles bearing these subject headings. Table 11.18 describes several title-keyword searches with such headings.

Table 11.18. Frequently-occurring Subject Headings in Title-keyword Searches

User query	Explanation
vietnam war	Blue System retrieves 8 titles. Of the 5 rated "very useful" by the user, most bear subdivided forms of the main subject heading, "Vietnamese Conflict, 1961–1975."
american events	Blue System retrieves 7 titles. The 1 title rated "very useful" bears the subdivided subject heading "United States — History — Chronology." Examples of titles retrieved with this subject heading are *The encyclopedia of American facts and dates* and *Webster's guide to American history: a chronological, geographical, and biographical survey*.
computing dictionary	Blue System retrieves 2 titles. 1 of the 2 titles rated "very useful" bears the subdivided subject heading "Computers — Dictionaries." Examples of additional titles retrieved with this subject heading are *The computer dictionary*, *Computer dictionary*, and *McGraw-Hill dictionary of electronics and computers*.
turbomachinery	Blue System retrieves 8 titles. The 3 titles rated "very useful" bear the subject heading "Turbomachines." A title retrieved with this subject heading is *Theory of turbomachines*.

Perseverance was not as much of a problem in title-keyword searches as it was in controlled vocabulary searches. Eight title-keyword searches were marred by perseverance. Despite the low postings (one or two titles) in four of the eight searches, users walked away from these searches without giving relevance assessments to retrieved titles.

The failure of five title-keyword searches was attributed to database failure. Most of these searches retrieved a few titles but they were false drops. An example was the title-keyword search for "workstation" that retrieved one title entitled *Workstation environment for wastewater treatment design*.

11.5 Keyword in Subject Heading Field Searches

Only the Blue System featured keyword in subject heading field searches. The Blue System submitted queries exceeding one word to keyword in subject heading field searches following unsuccessful exact, alphabetical, keyword-in-main-heading, keyword-in-subdivided-heading searches, and title-keyword searches or in response to the user's selection of the expand option following successful controlled vocabulary and/or other free-text searches. There were only four keyword in subject heading field searches. Two searches each were conducted in Earlham's catalog and in UM-D's catalog. The small numbers of keyword in subject heading field searches was due to the several searches preceding this search to which the search trees submitted user queries.

Three of the four searches were simple, successful searches in which users retrieved a few titles and rated them "very useful." The fourth search failed because of a vocabulary problem. The query "parallel programming languages" retrieved one title that the user rated "possibly useful." If the "languages" element of the query were deleted from this query, the Blue System could have retrieved several additional titles on the user's topic of interest.

Precision scores for the four searches were not very high. The two searches conducted in UM-D's catalog averaged 50.0% precision and the two searches conducted in Earlham's catalog averaged 35.3% precision.

11.6 Keyword-in-record Searches

11.6.1 Review of Keyword-in-record Searches

Both Blue and Pinstripe Systems featured keyword-in-record searches. The Blue System submitted queries exceeding one word to keyword-in-record searches as a last resort, that is, following unsuccessful exact, alphabetical, keyword-in-main-heading, keyword-in-subdivided-heading searches, title-keyword, and keyword in subject heading fields searches or in response to the user's selection of the expand option following successful search approaches. A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The Pinstripe System responded with this search regardless of the particular elements in user queries.

The two systems retrieved titles that contained the same words in title, subject heading, and other subject-bearing fields of bibliographic records as in user queries and displayed the first of one or more retrieved records to searchers.

11.6.2 Precision in Keyword-in-record Searches

Figure 11.5 shows precision scores for keyword-in-record searches. (Precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

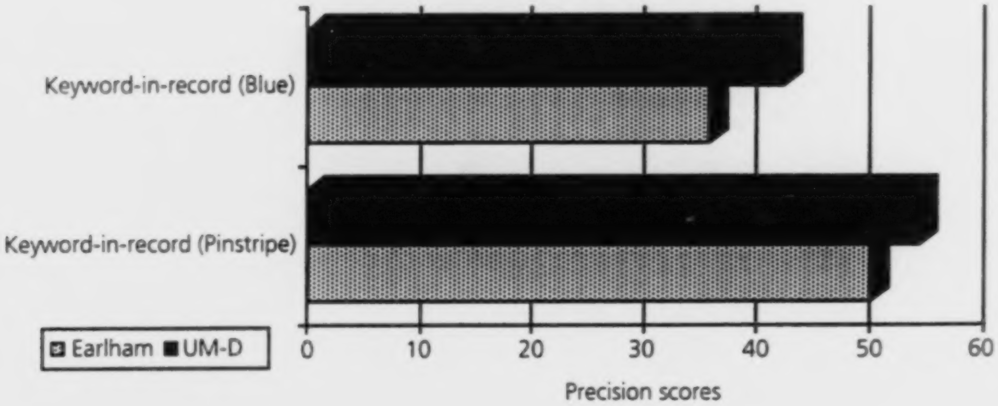


Figure 11.5. Precision in keyword-in-record searches

Precision scores for keyword-in-record searches were higher in the Pinstripe System than in the Blue System. However, precision scores in the Blue System were based on the results of about a half dozen searches.

11.6.3 Successful and Unsuccessful Keyword-in-record Searches

Table 11.19 summarizes the number of keyword-in-record searches and reasons for success or failure.

Table 11.19. Keyword-in-record Searches

Reasons for success/failure	UM-D		Earlham	
	Blue	**Pinstripe**	**Blue**	**Pinstripe**
Successful searches:				
Simple search	1 (16.7%)	93 (52.0%)	5 (100%)	19 (44.2%)
Unsuccessful searches:				
Perseverance	1	6	0	4
Too many retrievals	0	17	0	3
Database failure	1	40	0	7
Query specificity	0	6	0	3
Vocabulary	3	9	0	7
Spelling problems	0	2	0	0
Search approach failures	0	6	0	0
Subtotal	5 (83.3%)	86 (48.0%)	0 (0.0%)	24 (55.8%)
Total	6 (100%)	179 (100%)	5 (100%)	43 (100%)

Like keyword-in-subdivided-heading searches, keyword-in-record searches were not very common in the Blue System. This was due to the several searches preceding this search to which the search trees had submitted user queries and, probably, had been successful making matches. Of course, users could always expand their searches to retrieve the results of keyword-in-record searches. Keyword-in-record searches were much more common in the Pinstripe System. The keyword-in-record search was one of three subject searching approaches in this system.

About half of keyword-in-record searches were successful. Keyword-in-record searches were low posted in the Blue System where they averaged 2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Searches ranged from zero to 2,324 titles. If postings from the twenty searches in which users terminated their searches without rating even the first title displayed were disregarded, keyword-in-record searches averaged 44.7 titles per search. This dramatic reduction in the average number of titles was due to several searches among the twenty searches that retrieved so many titles that users did not bother to display them.

On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. Users who want additional material on their topics of interest could benefit from the system's retrieval of titles bearing these subject headings. Table 11.20 describes several keyword-in-record searches with such headings.

Table 11.20 Frequently-occurring Subject Headings in Keyword-in-record Searches

User query	Explanation
southern economy	Blue System retrieves 2 titles. User rates one title "very useful" bearing the subdivided subject heading "Southern states — Economic conditions."
slavery and black women	Blue System retrieves 2 titles. User rates one title "very useful" bearing the subdivided heading "Afro-American women — History — 19th century." Examples of titles retrieved with this subject heading are *Collected Black women's narratives* and *We are your sisters: Black women in the 19th century.*
black families	Pinstripe System retrieves 19 titles. User rates 9 titles "very useful;" all bear the subject heading "Afro-American families."
automatic transmission	Pinstripe System retrieves 7 titles. User rates 1 title "very useful" bearing the subdivided subject heading "Automobiles — Transmission devices."
code	Pinstripe System retrieves 16 titles. User rates 3 titles "very useful;" all bear the subject heading "Pulse-code modulation."

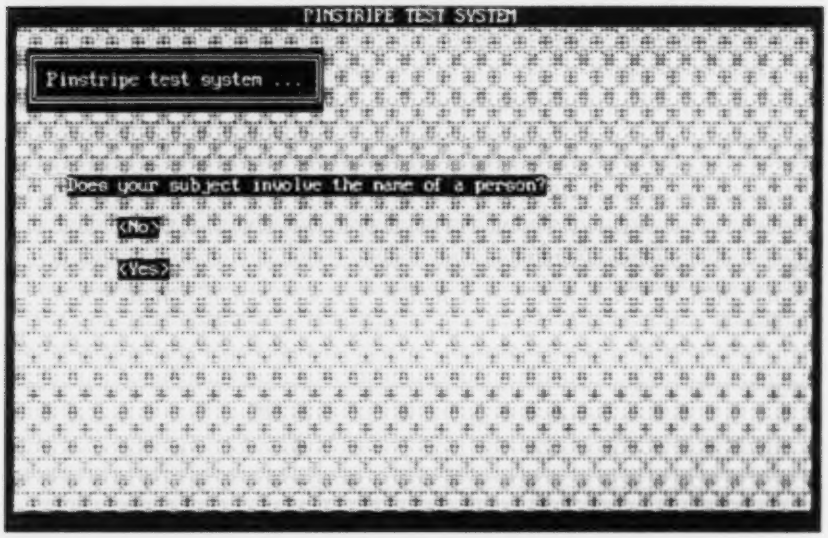
Perseverance was not as much of a problem in keyword-in-record searches as it was in controlled vocabulary searches. Eleven keyword-in-record searches (4.7%) were marred by perseverance. About twice as many (20) searches were marred by a different type of perseverance — too many retrievals. These searches averaged 411 titles; they ranged from 61 to 2,324 titles. Users did not rate the first title that the systems automatically displayed in these high-posted searches nor did they complete post-search questionnaires. They just walked away.

The failure of 48 (20.6%) keyword-in-record searches was attributed to database failure. Most of these searches retrieved a few titles but they were false drops. Examples are keyword-in-record searches for "propeller" that retrieved one title entitled *Measuring and metering of unsteady flows* containing the word "propeller" in a contents note, i.e., "Responses of small propeller current meters to impulsive and pulsating turbulent flows," and for "black church history" that retrieved one title entitled *Quest for equality: the life and writings of Mary Eliza Church Terrell, 1863–1954* bearing the words "black" and "women" in the series title (i.e., "Black women in United States history") and the word "church" in the title.

11.7 Subject Searches for Personal Names

11.7.1 Review of Personal-name Searching

Prior to allowing users to enter subject queries, the experimental online catalogs asked users whether their queries involved a person's name. The catalogs asked this question in the same way in the Blue and Pinstripe Systems (see figure 11.6). The catalogs responded to users who gave positive answers to this question with a series of pop-up screens that prompted users to enter surname, given name, and topic (see figures 5.29–5.31).



PINSTRIPE TEST SYSTEM

Pinstripe test system ...

Does your subject involve the name of a person?

Figure 11.6. Asking users whether their queries involve names

End users entered subject searches for personal names into the experimental online catalogs at UM-Deaborn and Earlham College. The database for the former was composed of library materials in computer science and technology. Few personal names were assigned to UM-D bibliographic records. In fact, the entire alphabetical list contained less than 150 personal-name subject headings.

The database for Earlham College was composed of library materials in American history. Many names were assigned to bibliographic records. In fact, the alphabetical list of personal-name subject headings contained almost as many unique names as the alphabetical list of topical subject headings.

BEST COPY AVAILABLE

11.7.2 Queries for Personal Names

11.7.2.1 Unusable Queries at UM-D

Despite the infrequency with which personal names were connected with computer science and engineering topics in the UM-D database, some UM-D users responded positively to the system's question about personal names and entered queries bearing personal names. Examples were:

- limbrough [sic] (surname), rush (given name)
- pasteur (surname), louis (given name)
- regan [sic] (surname)
- corniliustacitus [sic] (surname)
- kevorkian (surname), physician assisted suicide (topic)
- perot (surname), ross (given name), petroleum (topic)
- jakeson [sic] (surname), michael (given name), singing muise [sic] (topic)
- ogata (surname), control theory (topic or title)

These eight queries were typical of the personal-name queries for subjects that users entered into the experimental online catalog at UM-D. Most queries were out of scope (first six queries listed above). A few queries were characterized by playing around (seventh-listed query above) or author-title searches (last-listed query above).

Two subject queries for personal names that UM-D users entered into the experimental online catalogs deserve mention, "pascal" and "gauss." The former is a name of a computer programming language or a unit of pressure. The latter is an electromagnetic unit of magnetic induction. Both queries were valid ***subject*** queries. The systems submitted both queries to subject searching approaches for personal names because users responded positively to the system question about personal names being involved with queries. In fact, both queries did involve names of persons, Blaise Pascal, a seventeenth-century French scientist, and Johann Karl Gauss, a nineteenth-century German mathematician.

The question about names in queries that users answered prior to entering their queries was quite effective in categorizing subject queries as personal names or subject queries generally. In view of the two queries discussed above (i.e., "pascal" and "gauss"), it was not entirely effective in effecting a correct categorization. Future computer systems that separate subject searches for personal names or for subject queries generally could experiment with different wording for the question that categorizes subject queries as personal names or subject queries generally.

A handful of UM-D users answered negatively to the question about names in queries. They entered personal names (and, sometimes, title words) into the catalogs. Examples were:

- assad
- hitler
- bush, g
- sinclair upton
- schaum's
- schaum's outline series dynamics
- dijkstra

Some names were out of scope (first four names listed above). Other names were probably attempts at author or author-title searches in the areas of computer science and engineering (last three names listed above). Why did users answer negatively to the question about personal names that resulted in the system's submission of these names to subject searching capabilities for subjects generally? In the absence of post-search interview data, we can only speculate on such reasons. Some users could have changed their minds and entered different queries than the ones they originally intended to enter. Other users could have misunderstood the question about personal names. Still other users could have failed to read the question about personal names.

Future computer systems that separate subject searches for personal names or for subject queries should experiment different wording for the question that categorizes subject queries as personal names or subject queries generally. Users who become frequent users of a particular online catalog could become bored with answering the same question over and over again. They might need a different method of submitting queries for personal names than users who use the system infrequently.

11.7.2.2 Unusable Queries at Earlham

Personal-name subject headings were quite common in the bibliographic database on American history at Earlham. Yet findings about the entry of queries for personal names were quite similar to those at UM-D. End users at Earlham entered queries for personal names that were out of scope. Some could have been author-title searches for known items. Examples were:

- calvin (surname), john (given name), treatises against the anabaptists (topic)
- chopin (surname), kate (given name)
- poe (surname), edgar (given name)
- schlesinger (surname), arthur (given name), the age of jackson (topic or title)
- arendt (surname), hannah (given name), totalitarianism (topic)
- webster (surname), dictionary (topic or title)

On occasion, end users at Earlham answered negatively to the system's question about personal names in queries. Consequently, the system submitted their queries to subject searching capabilities for subjects generally. Below are examples of out of scope queries bearing personal names that were submitted to such capabilities:

- tanner, henry
- marilyn monroe
- rodin

11.7.2.3 Elements in Usable Queries at Earlham

Users entered 27 different queries for personal names into the experimental online catalogs at Earlham. Some queries contained topical elements:

- jackson (surname), andrew (first name), jacksonian democracy (topic)
- calhoun (surname), john (given name), state rights (topic)
- sumner (surname), charles (given name), reconstruction (topic)

Other queries only contained name element(s):

- roosevelt (surname), eleanor (given name)
- rustin (surname), bayard (given name)
- fremont (surname)

Of the 27 queries for personal names, a few deserve special mention because of ways in which users stated name and topic elements. Both Blue and Pinstripe Systems prompted users for surname, given name, and topic elements. Although they did not prompt users for middle name elements, four queries contained two elements in the given name pop-up box that represented first and middle name elements. These queries were:

- king (surname), martin luther (first and middle names entered into first name pop-up box), about the bycotts [sic] they had way back in mississippi (topic)
- king (surname), coretta scott (first and middle names entered into first name pop-up box), learn about what happened in 1927 (topic)
- king (surname), martin luther (first and middle names entered into first name pop-up box), about in the 1925 (topic)
- washington (surname), booker t (first name and middle initial entered into first name pop-up box), african american history (topic)

The first three queries were entered successively into the experimental online catalog. They were probably entered by the same individual. In view of our experience of the four queries above, we could speculate that some users may feel compelled to enter the middle name or initial for famous people that are known by their full names. Yet, in the following query for an author who is known by her full name (i.e , Harriet Beecher Stowe), the user did not enter the middle name:

- stowe (surname), harriet (first name), uncle tom's cabin (topic)

Two other queries were especially interesting:

- hoover herbert (surname and given name entered into surname pop-up box), herbert (first name), economic (topic)
- robert lafollette (given name and surname entered into surname pop-up box), mcarthy (topic)

In the first-listed query above, the user entered both surname and given name in inverted form into the surname pop-up box. He repeated the entry of the given name in the given name pop-up box. Perhaps he realized he entered both first and last names into the surname pop-up box after the given name pop-up box prompted him to enter the given name. When the user was ready to submit the entire query to the system, the system gave the user the opportunity to change entered name and topic elements but he did not make changes. The Pinstripe System performed a keyword-in-record search for name and topic elements in this query and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements but it continued searching by placing the user at the personal-name subject heading in an alphabetical list of personal-name subject headings beginning with the name "Hoover, Herbert, 1874–1964." Unfortunately, this Blue-system search was marred by user perseverance.

In the second query, the user gave the name in direct form. The Pinstripe System performed a keyword-in-record search for name and topic elements and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements but it continued searching by displaying to the user an alphabetical list of personal-name subject headings that began with the letter "r." The user browsed backward repeatedly and found the subject heading "LaFollette, Robert Marion, 1855–1925." She selected this subject heading, displayed eight titles, and gave two titles "very useful" ratings.

The ASTUTE experimental online catalog was not unique in its division of subject searches into subject searches for personal names and for subjects generally. UCLA's ORION online catalog also enlists such a division. The ASTUTE experimental online catalog was unique in its prompting of users for the name and topic elements of personal-name queries. Additional research is needed to determine how users enter a wide range of names, especially names in which the distinctions among surname, given-name, and middle-name elements are not as clear as the names entered by Earlham College users, for example, foreign names (Vincent Van Gogh, Jan van Eyck, Rogier van der Weyden) Native American names (Tecumseh, Blue Jacket, Geronimo), Classical names (Hippocrates, Aristotle), and to determine whether systems can provide satisfactory results.

11.7.2.4 Subject Searching Approaches for Personal-name Queries

When users entered subject queries for personal names, the Blue System was governed in its selection of a subject searching approach by the search trees for personal-name queries (see Table 4.1). The system first submitted queries with name and topic elements to the keyword-in-record search. If it failed to retrieve titles bearing these elements, it ignored the topic element and responded with the results of an alphabetical search in which listed subject headings for personal names were in the same alphabetical neighborhood as the last and first name elements of the query. The Pinstripe System always responded to personal-name queries with the keyword-in-record search.

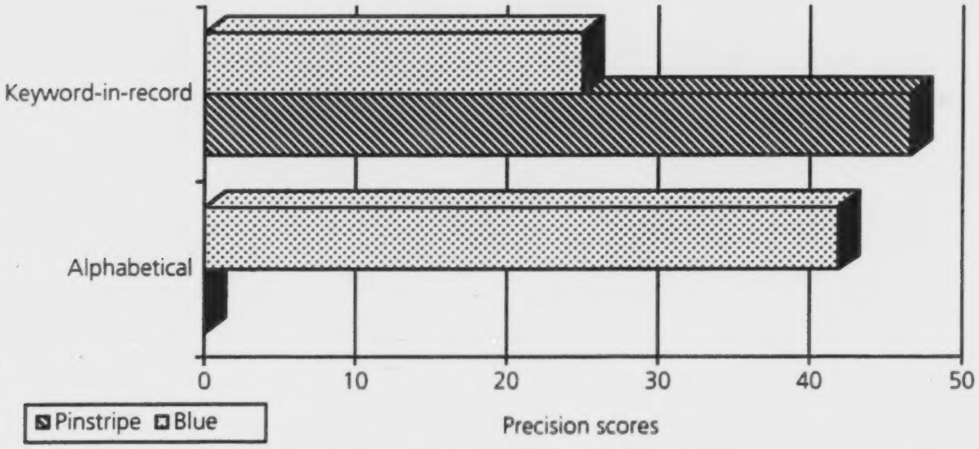


Figure 11.7. Precision in personal-name subject searches

Figure 11.7 shows precision scores for personal-name subject searches. (Precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

Precision scores for keyword-in-record searches were higher in the Pinstripe System than in the Blue System. Precision scores for the Blue System's alphabetical approach and the Pinstripe System's keyword-in-record search were about the same.

Table 11.21 summarizes the subject searching capabilities that handled subject queries for personal names in the Blue and Pinstripe Systems and characterizes them as successful or unsuccessful.

Table 11.21. Subject Searches for Personal Names

Reasons for success/failure	Earlham	
	Blue	Pinstripe
Successful searches:		
Keyword-in-record	4	8
Alphabetical	8	n/a
Subtotal	12 (52.2%)	8 (38.1%)
Unsuccessful searches:		
Perseverance	6	0
Database failure	4	4
Search approach failures	0	9
Query specificity	1	0
Subtotal	11 (47.8%)	13 (61.9%)
Total	23 (100%)	21 (100%)

Over half (12 of 23) of subject searches for personal names were successful in Earlham's Blue System. About a third (8 of 21) of subject searches for personal names were successful in Earlham's Pinstripe System. The most frequent reasons for failure in the Blue and Pinstripe Systems were perseverance and search approach failures, respectively. Both systems had their share of database failures.

Of the 27 subject queries for personal names, all but four queries featured elements for names and topics. The keyword-in-record search failed to retrieve titles for half of these queries because of the topic element. Frequently, this resulted in a search approach failure for the Pinstripe System (see section 10.3.1.6). For example, a user entered the query "sumner" (surname), "charles" (first name), "reconstruction" (topic) into the experimental online catalog. The Pinstripe System submitted this query to a keyword-in-record search and failed to retrieve titles. The Blue System also submitted this query to a keyword-in-record search. When it failed to retrieve titles, the system ignored the topic element and responded with the results of the alphabetical approach. The user selected the personal-name subject heading "Sumner, Charles, 1811–1874" from the alphabetical list, retrieved eight titles, and rated all eight titles "very useful."

Search approach failures were quite common for the personal-name queries bearing topic elements that users entered into the Pinstripe System. The failure of nine subject queries was attributed search approach failures in the Pinstripe System.

Several Blue-system searches were successful in retrieving titles through the alphabetical approach to which users gave positive relevance assessments even though the titles did not specifically refer to the topic element of their queries. The failure of a few Blue-system searches to retrieve useful titles was attributed to user perseverance, query specificity, or database failure. For example, after the query "whittier" (surname), "john" (given name), and "abolition" (topic) failed to retrieve titles in a keyword-in-record search, the Blue System responded with the alphabetical approach but the alphabetical list of personal-name subject headings did not contain a subject heading for John Whittier. The failure of this Blue-system search was attributed to database failure.

On occasion, users were not satisfied with the results of Blue-system searches in which the alphabetical approach ignored the topical element of their queries and retrieved titles based only on the name elements of their queries. For example, in response to the query "king" (surname), "coretta scott" (first name), "learn about what happened in 1927" (topic), the Blue System ignored the topical element because of its inability to retrieve titles through a keyword-in-record search and responded with an alphabetical list of personal-name subject headings. The user selected the personal-name heading for Coretta Scott King, retrieved several titles, and gave them all negative relevance assessments. Perhaps this user would have only been satisfied with titles that specifically referred to Mrs. King in 1927.

11.7.2.5 Recommendations for Personal-name Queries

The Blue System performed much better than the Pinstripe System in terms of responding with a subject searching approach that led users to titles on their topic of interest or on a facet of their interests. Search approach failure occurred in the Blue System with much less frequency; however, the failure of a few Blue-system searches was attributed to user perseverance, query specificity, or database failure.

We recommend the implementation of the search trees for subject queries for personal names in operational online catalogs. The search trees should be accompanied by pop-up windows that prompt users for the elements of their personal-name queries. This prompting enables systems to differentiate between name and topic elements of queries and choose to ignore certain elements in their attempts to retrieve titles. More research is needed to ensure that users enter the correct elements in various pop-up windows but this should not deter system designers from implementing the search trees and prompting mechanisms in operational systems.

The procedures for prompting users for personal names in subject searches should be extended to prompting users for personal names in author searches and for personal names and titles in a combination author-title search. This could help eliminate the frequency with which users enter direct forms of names in author searches into systems that expect indirect forms of names in author searches and respond with an alphabetical list of author names based on the left-most name element in user queries. It could also help systems make informed decisions about the particular element(s) they should ignore in their attempts to retrieve titles. For example, if systems fail to retrieve titles using both author name and title elements of author-title queries, they could ignore the author or title elements of queries in their attempts to retrieve titles.

11.8 Librarians' Subject Searching Experiences

11.8.1 Introduction

Eight librarians participated in the Comparison Search Experiment. One librarian was a cataloger and seven librarians were reference librarians. They accounted for a total of thirty-three search administrations. The analysis of subject searches in chapters 7–11 intermingled searches performed by end users and searches performed by librarians. We did not separate results for librarians from results for end users generally.

Comparison Search Experiments with librarians included personal interviews following their searches in the Blue and Pinstripe Systems. We placed their remarks into the following broad categories: (a) subject searching capabilities, (b) number of retrievals, (c) expand capability, (d) browsing, (e) search trees, (f) search administration problems, (g) database, (h) personal-name searching, and (i) housekeeping features. Most comments were paraphrased versions of the comments librarians made. Comments placed in quotation marks were actual statements made by librarians.

11.8.2 Categorized Comments

a. Subject searching capabilities

a1. General comments

- I am waiting for the system to tell me what types of subject searches I can do. I like search menus like Mirlyn has with searches for s= (subject heading browse search) and k= (keyword search with implicit Boolean).
- When the system finds only a few titles for your query, it should have capabilities for finding more titles. (2 librarians)
- System should have been able to ignore the "action" word in my query, "origins of the ku klux klan," and retrieved titles for the "ku klux klan" component of the query.

- Following a keyword-in-record search with few retrievals, the system should have suggested more subject searching terms to help me find more titles. (2 librarians)

a2. Boolean searching

- It is important for me to know if I can search for more than one term at a time or if I can combine terms.
- The system could not handle a Boolean query, i.e., (("blacks or afroamericans or chinese) and west)." Add a Boolean searching capability because some of our students know how to construct Boolean search statements.
- I am unsure whether the system will allow explicit truncation.

b. Number of retrievals

- Our library's online catalog shows you individual postings for your query words so you have an idea which word is detrimental to your query.
- Add the number of retrievals for listed subject headings so I know whether I want to limit my search.

c. Expand capability

- I like the expand function because it helped me find more titles on the topic I wanted. (4 librarians)
- I expect that the "expand" capability finds more titles using broader and narrower terms. (3 librarians)
- By choosing "expand," I had hoped to retrieve a list of broader and narrower terms. Instead, the system responded with the same titles that I saw previously.
- I like the expand function. It didn't help in my search today but I might need it in the future. (2 librarians)
- "Expand" capability seems like a "black box." I sure would like to know what it is doing.

d. Browsing

- I liked the keyword-in-subdivided-heading list because listed terms allowed me to narrow my topic or find a more precise term for my topic of interest.
- I liked the option to browse broader and narrower terms.

d1. Alphabetical lists

- When my search fails to retrieve titles, the catalog should show me a list of subject headings to help me find relevant searching terminology. (4 librarians)
- I liked when the system responded with an alphabetical list of subject headings; however, I was perplexed when it didn't respond with this list. I think searchers would find this inconsistency perplexing also.
- I preferred the Pinstripe System because it responded to my query "human powered vehicles" and "virtual reality" with an alphabetical list of terms. I found the terms "Human locomotion" and "Virtual computer systems" which are somewhat close to my interests. The Blue System merely told me it could not find titles on my topic of interest. (2 librarians)
- Neither system found titles for my query "computer-assisted instruction. " At least they could have responded with an alphabetical list of terms.
- I like to look at an alphabetical list of subject headings to find more titles. I wish this system had such a capability. (2 librarians)

d2. Broad categories

- Blue System makes me go through too many levels of selecting subdivided terms before I am able to display titles. In my search for "civil rights," I had to select four or five levels of terms before I could display titles assigned to the subject headings "Civil rights — United States — History — 20th century" and "Civil rights demonstrations — Florida — St. Augustine — History — 20th century." And then I only retrieved one or two titles! If I'm only going to retrieve two titles, can't I display these titles sooner?
- I would like to see subdivided and unsubdivided subject headings in a single list. The other approach (broad categories) is too complicated for selecting subdivided subject headings.
- If searchers do not know LCSH, the broad categories list could help them.

e. Search trees

- "As a librarian, I want to know what the system is doing or I want to tell it what to do." The system should give me a clue as to what search to do next to produce retrievals.
- It bothers me that I don't know what fields the system is searching. I will sometimes do a title search and scan subject headings in retrieved titles to find a correct subject heading but I cannot perform this strategy because I

do not know how the system searches. "I like to be in control when I'm searching."

- I would like the system to give me a message telling me what it is doing. I don't know if this would be as much help to students as to librarians but students do pick up on what systems are doing.

f. Search administration problems

- I wanted to display titles but I must have hit the wrong key because the system switched systems on me.

g. Database

- I like the idea of having a system dedicated to computer science and technology because it is easy to block out extraneous material.

h. Personal-name searching

- In my search for "willke" (surname), "wendell" (given name), "nomination" (topic), the system retrieved general titles on Wendell Willke but I would have preferred titles specifically on his nomination.
- In my search for "jefferson" (surname), "thomas" (given name), "inventions" (topic), the system retrieved general titles on Thomas Jefferson but I would have preferred titles specifically on his inventions.

i. Housekeeping features

- I like the system's displays of retrieved titles (2 librarians)
- I like how the system blinks "Searching ..." as long as it is still searching for your terms.
- I like having both keyboard and mouse so I can choose between them.
- I like how the system gives me a chance to edit my query before I enter it.
- I like how the system suggests that my query might be misspelled. (2 librarians)
- I am not familiar with how the system indicates disabled options. Perhaps it would be better to limit displayed options to those that I can select.
- I don't like how you have to hit the <Enter> key twice to enter or select a search term.
- I like the windowing capabilities.
- I like how few keystrokes I have to enter to perform a search.

11.8.3 Suggested System Improvements

Librarians' comments call for the following improvements to systems that feature search trees:

- A capability for "set detail on." Such a capability would be available to librarians and end users who want to know how the system is processing their queries. This capability could reduce the mystery of both the search trees and expand capability because both submit user queries to various subject searches but do not report unsuccessful results.
- User feedback capability. User feedback could enlist several approaches to retrieving additional titles, especially following low-posted keyword searches. For example, systems could perform searches for subject headings common to the many titles that users retrieve and give "very useful" ratings.
- Best-match approach. The best-match approach would feature stemming, weighted-term probabilistic retrieval, and output ranking. Systems would perform searches on a best match, combinatorial basis. A match on fewer than all query-word stems may retrieve titles but titles that have all the query-word stems would be displayed first. This approach would have helped librarians retrieve titles for the query "origins of ku klux klan" query and Boolean query "((blacks or afroamericans or chinese) and west)." It would also address comments on finding additional titles following low-posted keyword searches.
- Rename the "expand" capability. When librarians became familiar with the operation of the "expand" capability, they found it valuable for retrieving additional titles on occasion. Some remarks, however, indicated that librarians thought this capability would retrieve lists of broader and narrower terms. This capability should be renamed to make it indicative of its function.
- Indicate titles retrieved and displayed previously. Blue- and Pinstripe-system users had several opportunities to retrieve and display the same titles repeatedly in subject searches. Both systems should indicate titles that users have retrieved in previous searches and displayed in previous searches. In a system that features icons, a check mark or an open eye could be placed adjacent to listed titles to indicate titles that users have retrieved and displayed in previous searches, respectively.
- Enlist the alphabetical approach as the search "of last resort." When the search trees fail to retrieve titles, systems should respond with the

alphabetical list of subject headings to give users ideas of terms in the alphabetical neighborhood of their entered term.

- Reduce the complexity of exact-approach displays. If exact-approach searches result in the retrieval of a few subdivided subject headings, systems should display these headings immediately instead of requiring users to proceed through several levels of broad categories to display subject headings and, ultimately, to retrieve titles.

Greater experience with the Blue System would make several aspects of the experimental system less uncommon or more familiar: (1) option disabling, (2) expanding, and (3) responding to pop-up menus to differentiate between more than one action following the entry or queries or selection of terms from lists.

11.9 Chapter Summary

Chapter 11 examined subject searching capabilities in the Blue and Pinstripe Systems. It focused primarily on the success of these capabilities but it also covered frequent reasons for their failure.

Controlled vocabulary searches were especially plentiful in the Blue System and accounted for almost two-thirds of searches conducted in the Blue System (section 11.2). About half (52.5%) of controlled vocabulary searches were successful in retrieving at least one title rated "very useful" by system users. Many (54.5%) of these successful searches were searches characterized by no browsing (see Table 11.5). That is, users displayed retrieved titles and terminated their searches without browsing available options for broader or narrower terms, subdivided terms, notes, or expanding the search. Only 14.4% of successful controlled vocabulary searches featured extensive browsing.

Frequent reasons for the failure of controlled vocabulary searches in the Blue System were perseverance and navigation problems (see sections 11.2.4 and 11.2.5). Some of the searches in which perseverance was a problem could have been navigation problems. Based on the analysis of searches that failed due to perseverance and navigation problems, the following suggestions were made for improving the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.

- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

The keyword-in-subdivided-heading search was a controlled vocabulary search but it was treated separately from other controlled vocabulary searches because it was independent of the exact approach (section 11.3). Keyword-in-subdivided-heading searches were much more frequent in the Pinstripe System than in the Blue System. The reason why only nine of 156 keyword-in-subdivided-heading searches were conducted in the Blue System was attributed to the search trees that submitted queries to the keyword-in-subdivided-heading search after submitting queries to three other controlled vocabulary searches (section 11.3.1).

Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing (see section 11.3.3). These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing.

Unsuccessful keyword-in-subdivided-heading searches outnumbered successful the keyword-in-subdivided-heading searches (see section 11.3.4). Almost three-quarters (112 of 156) of keyword-in-subdivided-heading searches were unsuccessful. Frequent reasons for failure were perseverance, database failure, and searches that retrieved no subject headings.

Title-keyword searches were unique to the Blue System (see section 11.4). Successful title-keyword searches outnumbered unsuccessful title-keyword searches. Almost 70% of title-keyword searches in UM-D's Blue System were successful; over 85% of title-keyword searches in Earlham's Blue System were successful. On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. When title-keyword searches retrieve few titles, users could find additional material on their topics of interest from the system's retrieval of titles bearing these subject headings.

Both Blue and Pinstripe Systems featured keyword-in-record searches but the majority (95.3%) of keyword-in-record searches were conducted in the Pinstripe System (section 11.6). About half of keyword-in-record searches were successful (see

Table 11.19). Keyword-in-record searches were low posted in the Blue System where they averaged 2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Large retrievals were the major problem with keyword-in-record searches in the Pinstripe System.

No subject queries for personal names that users entered into the experimental online catalog at UM-D were valid names (section 11.7). Most were out-of-scope or playing around. Users also entered subject queries for personal names that were deemed invalid into the catalog at Earlham; however, a total of 27 valid queries were entered. All but four of these queries contained topical elements. Blue-system search trees for handling personal-name queries were especially effective.

Post-search interview with librarians resulted in several suggestions for improving search trees. The most important suggestions were: (1) add a user feedback capability based on titles users rate "very useful," (2) add the best-match approach, (3) reduce the complexity of exact-approach displays, and (4) enlist the alphabetical approach as the search "of last resort."

The Blue System performed much better than the Pinstripe System in terms of responding with a subject searching approach that led users to titles on their topic of interest or on a facet of their interests. Search approach failure was typical of Pinstripe-system searches for personal names, especially when queries contained topic elements. The failure of a few Blue-system searches for personal names was attributed to user perseverance, query specificity, or database failure.

12 Redesigning Search Trees for Responsive Subject Searching

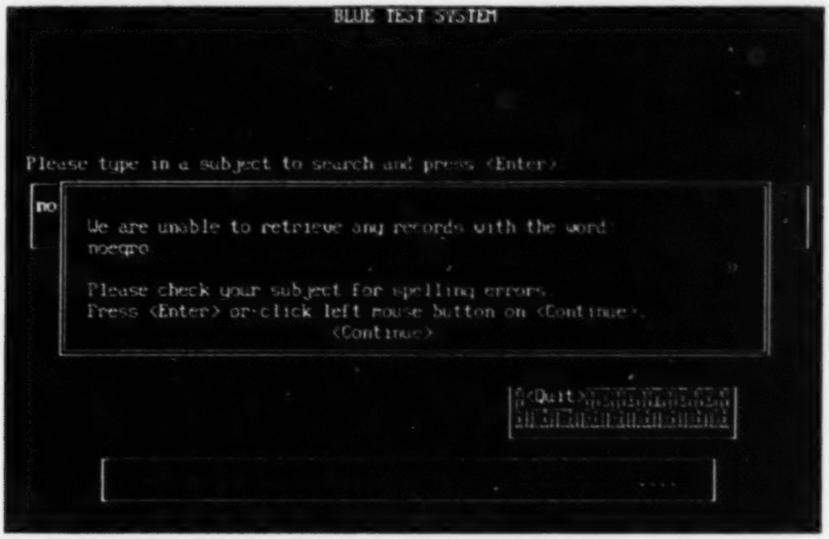
12.1 Introduction

Chapter 12 focuses on searching problems that have significant impact on the redesign of search trees — spelling and vocabulary problems. The bulk of the chapter is given to reconfiguring the search trees based on findings from the online retrieval tests that previous chapters feature (chapters 7–11).

12.2 The Need for Automatic Spelling Correction

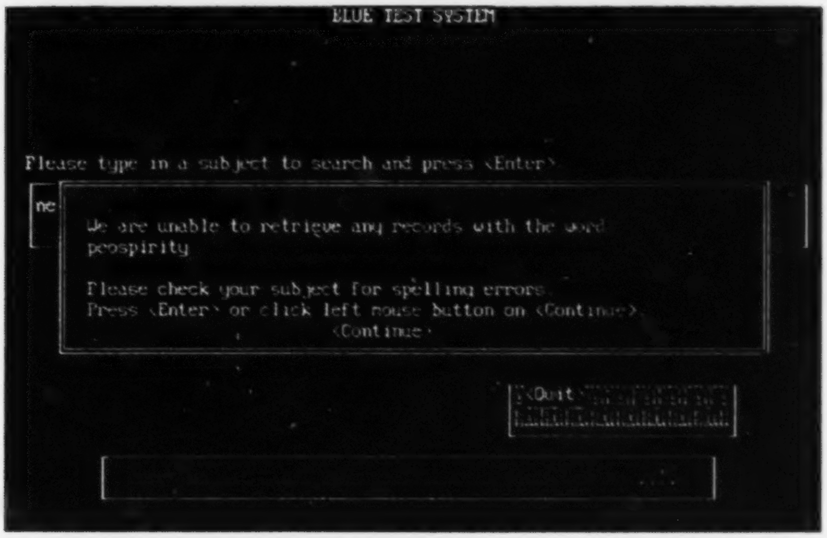
The ASTUTE experimental online catalog did not feature automatic spelling correction. It did, however, check user queries for possible spelling errors, inform users of query elements that failed to produce retrievals, and suggest users check such elements for spelling errors.

The Blue System checked user queries for subjects generally following the system's attempt to make a match in exact and alphabetical approaches. It compared words in user queries with words in a keyword index to determine whether the individual words in queries were used in the database. It checked the words in queries from left to right. An example is the user query "noegro peospirity in late 1920's." The system failed to find the word "noegro," informed the user of its failure, and suggested the user check spelling (see figure 12.1). The user corrected this word and the new query he submitted to the system was "negro peospirity in late 1920's." The system failed to find the word "peospirity," informed the user of its failure, and suggested the user check spelling (see figure 12.2). After responding to several system prompts to check spelling, the user eventually, the entered a query that contained no spelling errors, viz. "negro prosperity in late 1920's."



```
Please type in a subject to search and press <Enter>.
no

We are unable to retrieve any records with the word
noegro.

Please check your subject for spelling errors.
Press <Enter> or click left mouse button on <Continue>.
<Continue>
```

Figure 12.1 Informing users of possible misspellings



```
Please type in a subject to search and press <Enter>.
nc

We are unable to retrigue any records with the word
peospirity

Please check your subject for spelling errors.
Press <Enter> or click left mouse button on <Continue>.
<Continue>
```

Figure 12.2. Checking queries from left to right for possible misspellings

When the Pinstripe System's random selection algorithm called for the alphabetical approach, the system made no attempt to find possible spelling errors. When it called for keyword-in-subdivided-heading or keyword-in-record searches, the system performed the same error-checking routine as the Blue System. That is, it checked a keyword index to determine whether the individual words in queries were used in the

BEST COPY AVAILABLE

database. It checked the words in queries from left to right, informed users of query elements that failed to produce retrievals, and suggested users check such elements for spelling errors.

The experimental online catalogs responded to 134 queries with the message in figures 12.1 and 12.2 that informed users that their queries contained possible spelling errors. Table 12.1 describes what users did next.

Table 12.1. User Actions Following System Message Regarding Possible Spelling Errors

	UM-D		Earlham	
User Actions	**No.**	**%**	**No.**	**%**
Enter query on different topic	34	27.2	1	11.1
Enter same query	27	21.6	0	0.0
Quit search	15	12.0	2	22.2
Correct spelling	11	8.8	5	55.6
Enter same query minus unposted word	12	9.6	0	0.0
Enter same query and add new word	11	8.8	0	0.0
Enter new query with same stem as previous query	7	5.6	0	0.0
Enter singular or plural form of previous query	7	5.6	0	0.0
Enter acronym or spell it out	1	0.8	1	11.1
Total	125	100.0	9	100.0

Large numbers of users entered queries on different topics. Examples of queries bearing unposted words and the queries users entered following the system's message informing them of a possible spelling error are listed in Table 12.2. Words in italics were the unposted words that the experimental online catalog displayed to users for their correction.

Table 12.2. Succeeding Queries on Different Topics

Possibly Misspelled Queries	Next Queries on Different Topic
internet	usenet
microcad *graphing*	database
radiator design	heat transfer
zirconia	chemistry
general *relativity*	computers
The *z8* microcomputer	zilog microcomputers

Following the system's message informing them of a possible spelling error, large numbers of users entered the same query one or more times. Examples of such queries are "internet," "hovercraft," "androids," "barcode," "reinforce concrete," and "nanotechnology." Perhaps such users reentered the same queries because they wanted to make absolutely sure that the system had no titles on these topics. Reentering such queries, users might have been saying to themselves, "There's got to be information on this topic in here somewhere."

Following the system's message informing them of a possible spelling error, fifteen users at UM-D and two users at Earlham quit searching.

A total of sixteen users corrected the misspellings in their queries. Table 12.3 gives examples of misspelled queries and their corrected forms. Words in italics were unposted words that the experimental online catalogs suggested to users were misspelled.

Table 12.3. Misspelled Queries

Possibly Misspelled Queries	Corrected Queries
human power vechicles	human power vehicles
communcations	communications
carsuspension and handling	car suspension and handling
chemistryy	chemistry
abolishionism	abolitionism
c programming lanuage	c programming language
monitring performance of telephone operators	monitoring performance of telephone operators

Of the total 134 queries in which the experimental online catalog detected unposted words, 28 queries actually contained misspelled words. Users corrected sixteen of these queries. Examples of queries that users did not correct were "elecctronics," "circuts," "assemlby language," "cobal languages," and "ei\\EIFFEL." In response to the system's message informing them of a possible misspelling, some users quit searching, other users entered different queries, and still other users added new words or deleted the possibly misspelled words from queries.

A total of 23 users responded to the system message about possible misspellings by entering the same query minus the unposted word or by adding a new word to the same query. Table 12.4 lists user queries that prompted ASTUTE to respond with the possible-misspellings message and subsequent user queries with deleted or added words. Words in italics were unposted words that the experimental online catalogs suggested to users were misspelled.

Table 12.4. Queries with Words Added or Deleted

Original queries	**Added or deleted words in subsequent queries**
rotation of axis	rotation of axis formula
probability and *statisics*	probability
barcode	2d barcode
equilizer filters	filters
graphing	microcad graphing
cache memory	memory
internet	internet network
three-dimentional dynamics	dynamics

A handful of queries were placed in remaining categories. Examples of succeeding queries that had the same stem as preceding queries were "automanual" and "auto," "encoding" and "encode," "architectual design" and "architecture." Examples of succeeding queries that were singular or plural forms of preceding queries were "florida keys" and "florida key," "air bag" and "air bags," and "rotation of axis formula" and "rotation of axes formula." Two queries contained acronyms, "IWW" and "International Workers of the World" and "cobal languages" and "common business oriented language."

Frequently, the experimental system informed end users that their entered terms might be misspelled when, in fact, their entered terms were not posted in the database.

12.3 Vocabulary Problems and Search Trees

12.3.1 Introduction

The experimental online catalog contained information on the topics that interested several users. However, users phrased their queries in certain ways that prevented the experimental online catalogs from making retrievals in response to their queries. When users entered queries with such phrasing, the two experimental systems responded in different ways. The Blue System exhausted the subject searching approaches on the search trees before responding to users that it was unable to find titles to satisfy their queries. The Pinstripe System performed the search using one of the three subject-searching approaches at its disposal before responding to users that it was unable to find titles to satisfy their queries. However, differences between the two systems in terms of subject-searching approaches were transparent to end users.

Chapter 10 (sections 10.2.1.6, 10.2.2.6, 10.3.1.5, and 10.3.2.6) discussed vocabulary problems from the perspective of failed searches in full or partial administrations of the search experiment at UM-D and Earlham. While this section features some of the search queries discussed in chapter 10, it focuses on how the search trees could be enhanced to handle queries with vocabulary problems.

12.3.2 Punctuation in Queries

A few queries failed to retrieve useful titles on users' topics of interest because of punctuation in user queries and/or controlled vocabulary terms. For example, the user query "computer aided instruction" failed to match the hyphenated subject heading "Computer-aided instruction" because we did not double post the hyphenated word in this subject heading as two words, i.e., "computer" and "aided," and one word, i.e., "computeraided."

The original specifications for the experimental online catalog called for double posting of hyphenated words (Drabenstott and Vizine-Goetz 1994, 242). Our failure to double post such words was an oversight in the design of the experimental online catalogs. This failure, however, brought to light problems with punctuation, especially hyphenated words. For example, if a term in a subject heading or subdivision was not hyphenated, the system would not generate two normalized forms of the subject heading, a form for the one-word variant and a form for the two-word variant, because the term contained no hyphen. On occasion, users might enter the term with a hyphen. Since the system did not double post the term, it would fail to make a match. For example, systems would normally generate only one form of the subject heading "Coffee growers," that is, the two-word normalized form "coffee growers." This subject heading has no *see* reference under "Coffee-growers," so systems would not double post this heading under "coffee growers" and "coffeegrowers." If a user entered the query "coffee-growers," systems would fail to make a match because it generated no index entry under the normalized form of this query, viz. "coffee growers."

Systems could anticipate this unfortunate situation by finding additional hyphenated forms of words and phrases in various fields of bibliographic records, e.g., title and series-title fields, chapter titles in contents fields. For example, if a title contained the hyphenated term "coffee-growers," systems could search for subject headings or subdivisions bearing these two adjacent words and double post one-word and two-word normalized forms of these subject headings in the subject heading index, i.e., "coffee growers" and "coffeegrowers."

Retrieval was adversely affected by our failure to omit punctuation from one other user query. The user query "c.a.s.e." contained four periods. It retrieved several titles in a keyword-in-record search but all titles contained long contents notes with the initials "c., " "a.," "s.," and "e." If the periods and intervening spaces were deleted, the Blue System would match the *see* reference "CASE" for "Computer-aided software design" which probably described what the user had in mind.

12.3.3 Stemming User Queries

A case could be made for the inclusion of automatic stemming in the experimental online catalogs. Examples of queries that would have benefited from automatic stemming are described in Table 12.5.

Table 12.5. Stemming Queries to Enhance Retrieval

User query	Explanation
mechanical statics	Blue System exhausts subject searching approaches on the search trees and responds to the user that it is unable to find titles to satisfy the query. Pinstripe System submits this query to a keyword-in-subdivided-heading search and makes no retrievals. If the systems automatically stem the word "mechanical," they retrieve titles like *Engineering mechanics: statics, Schaum's outline of theory and probabilities of engineering: mechanics, statics, and dynamics,* and *Engineering mechanics: statics and dynamics.*
serial communication	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If automatic stemming is applied to this query, the experimental system retrieves *Mastering serial communications* in a title-keyword search.
reinforce concrete	Blue System exhausts search trees without retrieving titles or subject headings. If this query is stemmed, the system retrieves several titles bearing the subject heading "Reinforced concrete."
bar coding	Blue System exhausts search trees without retrieving titles or subject headings. The alphabetical approach in the Pinstripe System places users in the alphabetical neighborhood of the subject heading "Bar codes." Also, stemming produces retrievals under this subject heading.

12.3.4 Best-match Approach

The best-match approach would be fairly successful in producing retrievals for user queries. Table 12.6 describes queries for which the best-match approach could yield retrievals. The best-match approach should also feature stemming to facilitate retrieval of variant forms of user query words.

Best-match algorithms typically give greater weight to low-posted words than to high-posted words. A new weighting system ought to be devised for library cataloging databases because subject content is represented in so few fields of bibliographic records, i.e., title and subject heading fields. This system should increase the weight given to frequently-occurring words in subject headings and subdivisions because matches on such words are likely to retrieve subject headings describing users' topics of interest or a facet of their topics of interest. For example, the stems "women" and "slave" occur more frequently in subject headings than the stems "their," "treatment," and "society." Table 12.6 lists three titles on the user's topic of interest that the former high-posted stems retrieve.

Table 12.6. Queries Suited to the Best-match Approach

User query	Explanation
flow through porous media	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If this query is reduced to "porous media," the experimental system retrieves *Fundamentals of transport phenomena in porous media* in a title-keyword search.
fundamental of data structures in c	Pinstripe System's keyword-in-subdivided-heading approach fails to produce retrievals. If this query is reduced to "data structures in c," the experimental system retrieves *Data structures: an advanced approach using C* in a title-keyword search.
tolerance and geometric inspection	Both systems fail to produce retrievals. If this query is reduced to the stems of "tolerance" and "geometric," several titles are retrieved including *Geometric distancing and tolerancing* and *Tolerancing: computation on geometric uncertainties.*
black civil rights during the reconstruction	Both systems fail to produce retrievals. If this query is reduced to the stems of "black," "civil," "rights," and "reconstruction," titles such as *The Civil War and Reconstruction: a history of Black people in America, 1830–1880* and *The White response to Black emancipation: second-class citizenship in the United States since Reconstruction* are retrieved.
women slaves and their treatment by society	Both systems fail to produce retrievals. The stems of "women," and "slaves" retrieve titles such as *Women and the family in a slave society, Incidents in the life of a slave girl,* and *Six women's slave narratives.*
hopi indian land dispute	Both systems fail to produce retrievals. The stems of "hopi," "land," and "dispute" retrieve titles such as *The wind won't know me: a history of the Navajo-Hopi land dispute* and *The second long walk: the Navajo-Hopi land dispute.*

12.3.5 Go/see List

Another subject searching improvement would be the inclusion of a "go/see list" in online catalogs. Walker and Jones (1987, 68) described terms on the experimental Okapi system's "go/see list" as terms which the system treated as having the same meaning at search time. Examples of such terms were noun/adjective pairs, e.g., "French" and "France," acronyms and their spelled-out forms, e.g., "tv" and "television," and common, irregular plurals, e.g., "woman" and "women." When Okapi parsed user queries to find stopwords, it also searched for terms on the go/see list. Terms were not limited to one-word terms but sometimes included several words, e.g., "third world," "lesser developed countries," and "developing countries."

The research that led to the development of the search trees suggested a go/see list for one-word/two-word variations of concepts and noun/adjective pairs (Drabenstott and Vizine-Goetz 1994, 177, 292); however, ASTUTE was not equipped with go/see list functionality. In this research project, several queries could have benefited from a go/see list bearing alternative terminology, e.g., "blacks," "afro-americans," and "african-americans," "automobiles" and "cars."

12.4 Redesigned Search Trees

The analysis of online retrieval test data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. The search trees required enhancements to enable them to respond with useful retrievals to especially difficult user queries, that is, those for which the existing set of search trees produced no retrievals. This section presents the enhanced search trees.

12.4.1 Initial Search Tree

The search trees tested in this research project were described in chapter 4. The initial search tree remains basically unchanged from the original initial search tree (see figure 4.1). The only change from the original design is returning users whose searches failed to produce retrievals to the question about personal names. Based on user responses to this question, the new design of the initial search tree (figure 12.3) lets users distinguish their new or revised queries for personal subjects from queries for topical subjects generally and, based on summary characteristics that systems determine about the latter types of user queries, dispatches them to a particular search tree that favors certain subject searching approaches over others.

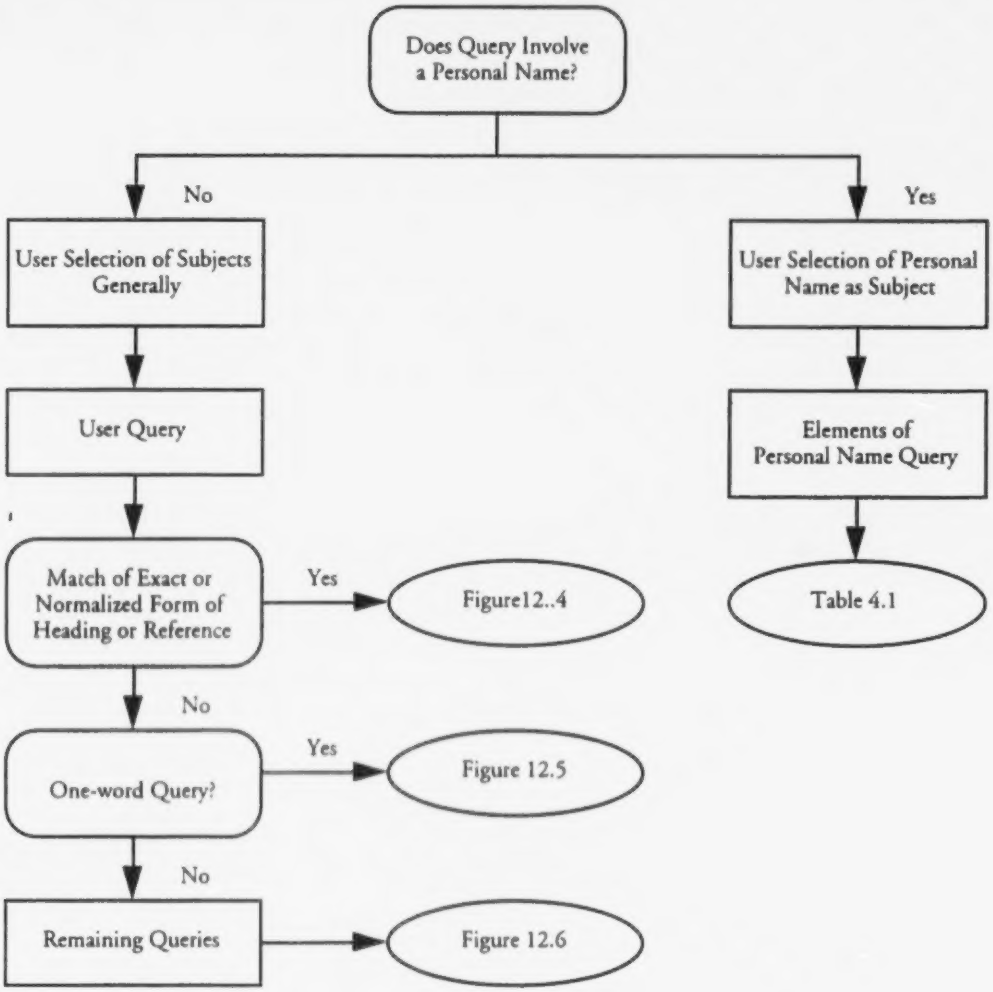

Does Query Involve a Personal Name?
No
Yes
User Selection of Subjects Generally
User Selection of Personal Name as Subject
User Query
Elements of Personal Name Query
Match of Exact or Normalized Form of Heading or Reference
Yes
Figure12..4
Table 4.1
No
One-word Query?
Yes
Figure 12.5
No
Remaining Queries
Figure 12.6

Fig. 12.3. Redesigned initial search tree

The search tree for the exact approach also remains unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach. Figures 12.4A and 12.4B show the search tree that features the exact approach.

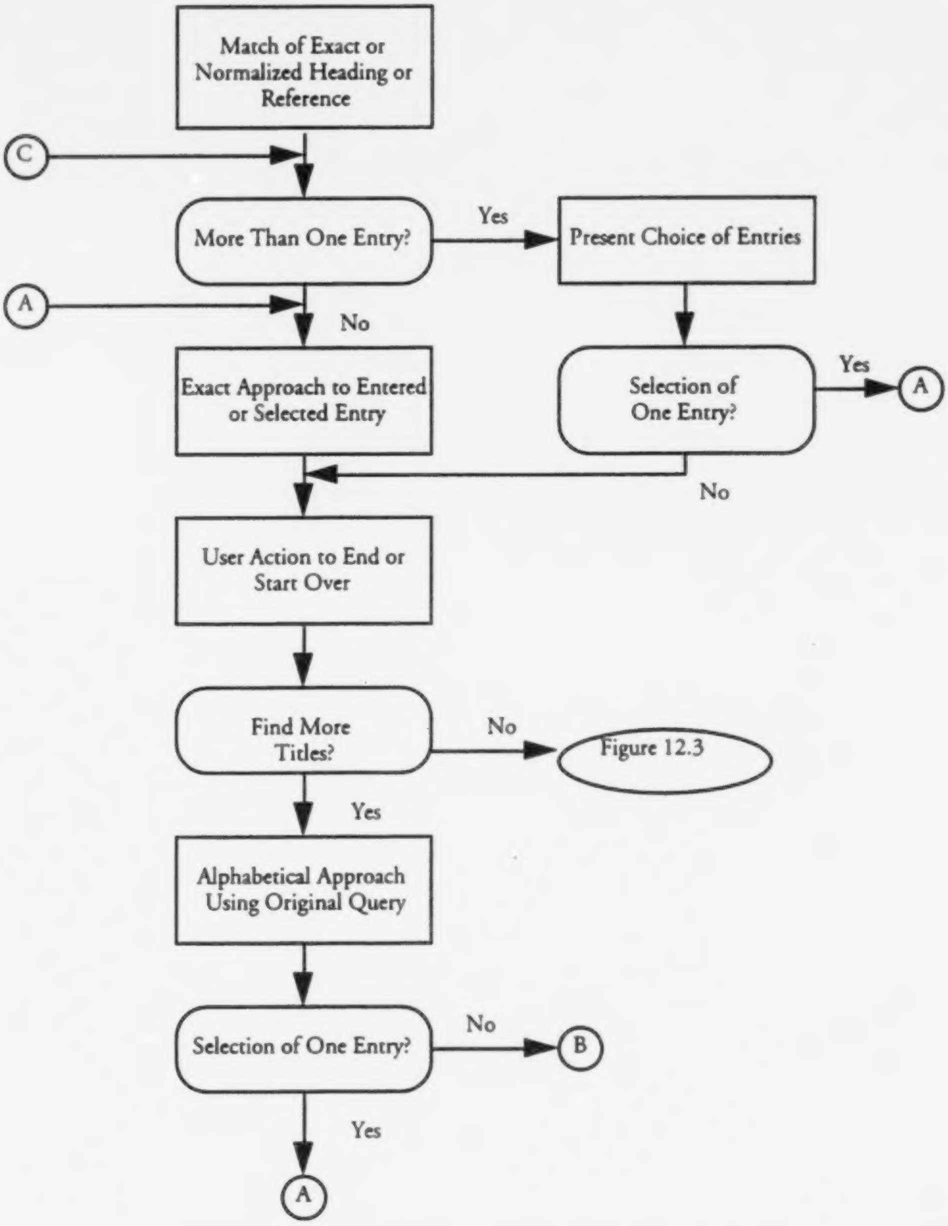

Match of Exact or Normalized Heading or Reference
C
More Than One Entry?
Yes
Present Choice of Entries
A
No
Exact Approach to Entered or Selected Entry
Selection of One Entry?
Yes
A
No
User Action to End or Start Over
Find More Titles?
No
Figure 12.3
Yes
Alphabetical Approach Using Original Query
Selection of One Entry?
No
B
Yes
A

Figure 12.4A. Search tree for the exact approach

In figure 12.4A, systems invoke the exact approach. This approach gives users options for browsing related terms, notes, subdivided forms of matched terms, and expanding the search. Users who want to expand their search are given the results of the alphabetical approach. Figure 12.4B shows the search approaches featured in search expansions. One-word queries are submitted to the keyword-in-main-heading search. Queries exceeding one word are submitted to the remaining controlled vocabulary searches, i.e., keyword-in-subdivided- and keyword-in-main-heading searches, and free-text searches, i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches.

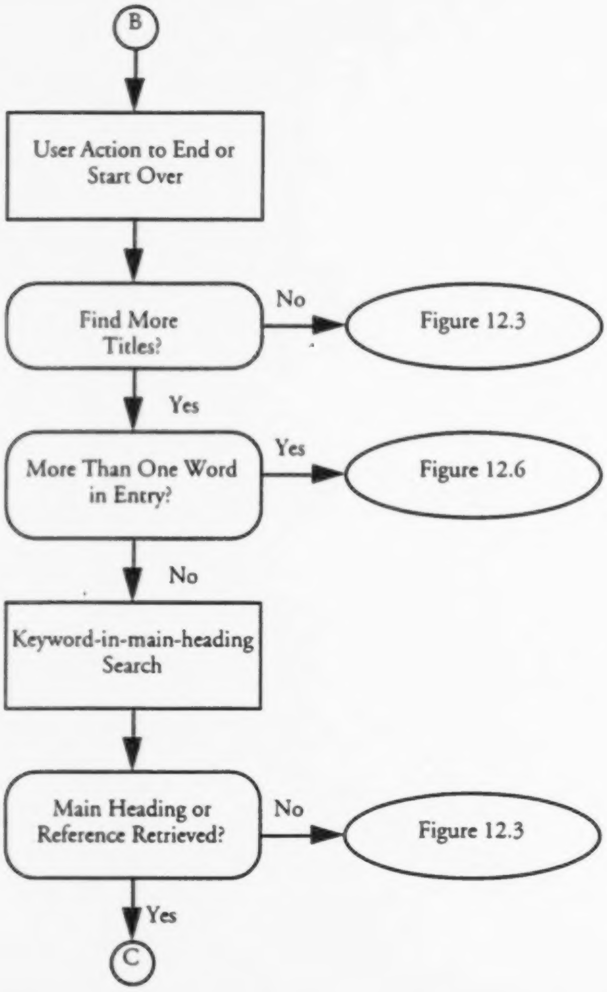

B
User Action to End or Start Over
Find More Titles?
No
Figure 12.3
Yes
More Than One Word in Entry?
Yes
Figure 12.6
No
Keyword-in-main-heading Search
Main Heading or Reference Retrieved?
No
Figure 12.3
Yes
C

Figure 12.4B. Search tree for the exact approach (contd.)

12.4.2 Search Tree for One-word Queries

The redesigned search tree for one-word queries contains a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names (figure 12.3), (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type "of last resort." The search tree for one-word queries is given in figures 12.5A and 12.5B.

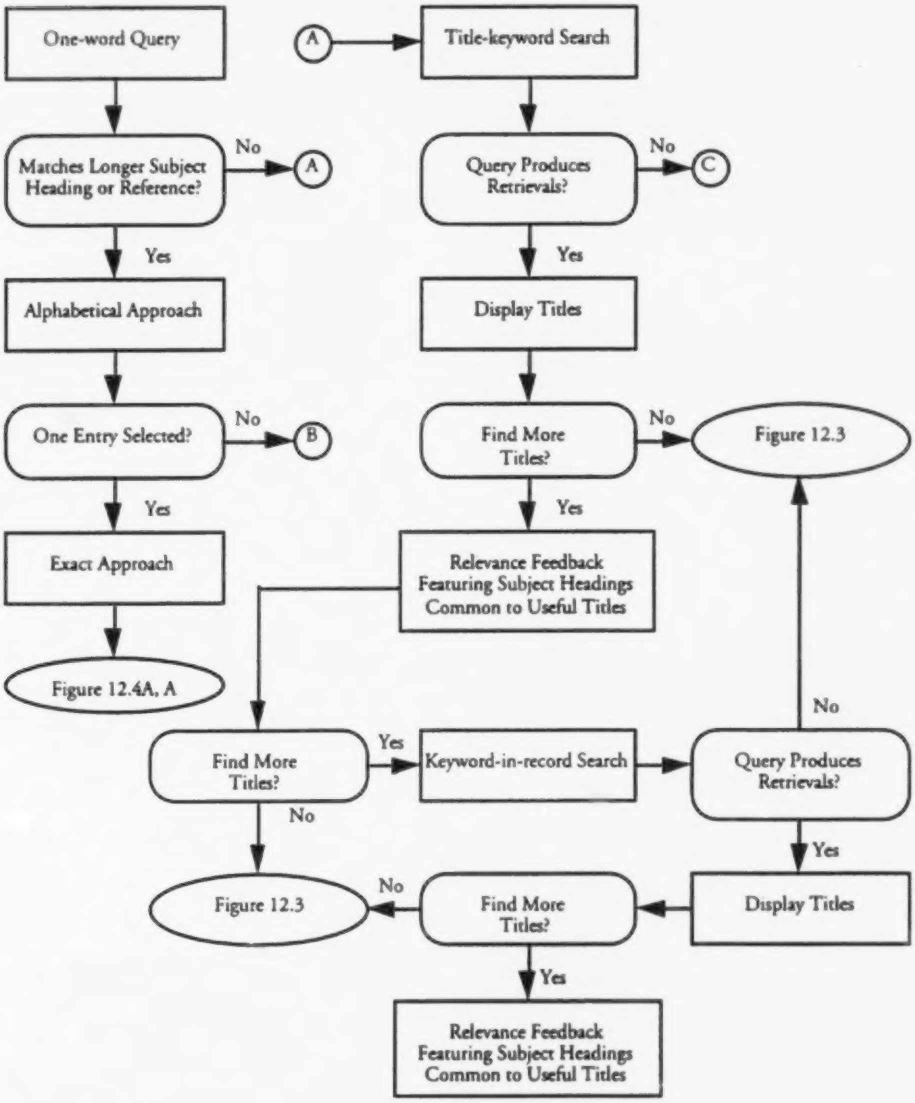

One-word Query
Matches Longer Subject Heading or Reference?
No
A
Yes
Alphabetical Approach
One Entry Selected?
No
B
Yes
Exact Approach
Figure 12.4A, A
A
Title-keyword Search
Query Produces Retrievals?
No
C
Yes
Display Titles
Find More Titles?
No
Figure 12.3
Yes
Relevance Feedback Featuring Subject Headings Common to Useful Titles
Find More Titles?
Yes
No
Keyword-in-record Search
Query Produces Retrievals?
No
Yes
Figure 12.3
No
Find More Titles?
Display Titles
Yes
Relevance Feedback Featuring Subject Headings Common to Useful Titles

Figure 12.5A. Redesigned search tree for one-word queries

Figure 12.5A shows system actions that test whether one-word queries meet the criteria for invoking the alphabetical approach. Queries that fail are submitted to title-keyword searches. If they are successful, systems retrieve titles, and, at the conclusion of the title display, they ask users whether they want to retrieve additional titles. Systems respond to users who give positive responses with the results of searches for subject headings common to several retrieved titles that users rated useful. Systems can continue to find additional titles through keyword-in-record searches and relevance feedback based on keyword-in-record search results. Of course, relevance feedback assumes that systems collect relevance assessments during the display of retrieved titles.

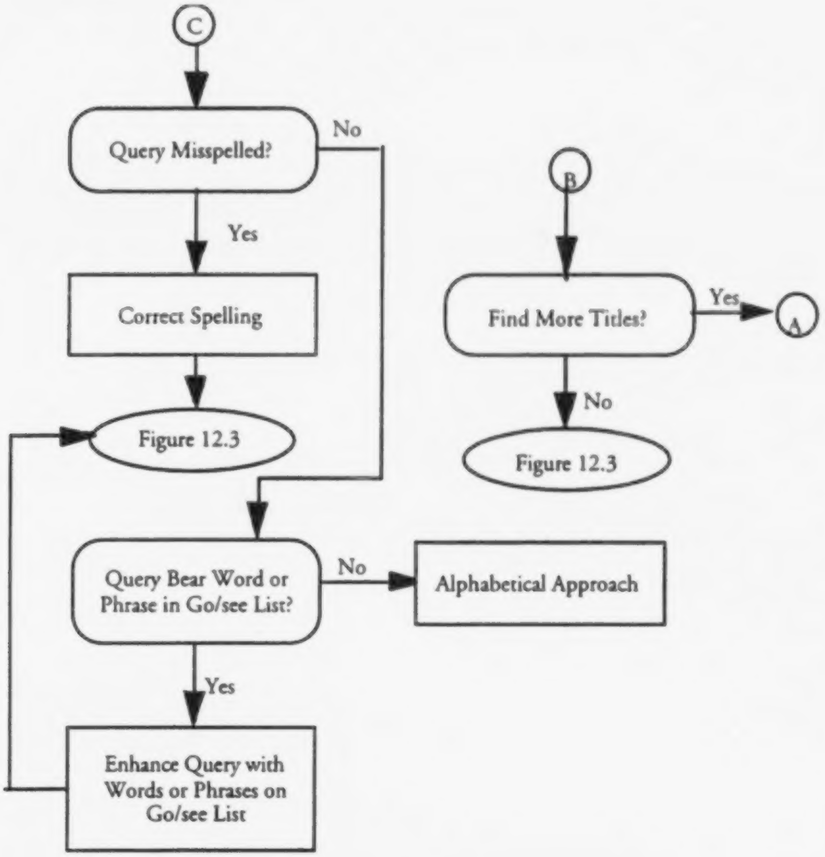

C
Query Misspelled?
No
Yes
Correct Spelling
Figure 12.3
Query Bear Word or Phrase in Go/see List?
No
Alphabetical Approach
Yes
Enhance Query with Words or Phrases on Go/see List
B
Find More Titles?
Yes
A
No
Figure 12.3

Figure 12.5B. Redesigned search tree for one-word queries (contd.)

The original search trees for one-word queries did not include keyword-in-record searches (see figure 4.3). This search type is included in the final version of the search trees for one-word queries. In view of the low levels of perseverance experimental online catalog users exhibited, few users will reach the results of keyword-in-record searches for one-word queries unless they continually expand search results. On occasion, however, keyword-in-record searches might be fruitful in retrieving useful titles for users whose queries fail to retrieve titles through exact, alphabetical, and title-keyword searches. For example, a contents note or summary field of a bibliographic record might bear a word matching the query. The only way for systems to retrieve this record is through a keyword-in-record search.

Figure 12.5B shows the operations that systems perform on queries that fail to produce retrievals in title-keyword searches. They first check user queries for possible misspellings. Systems could assist users in correcting spelling errors using spelling-correction routines similar to such routines for word processing programs in which they suggest alternate spellings for misspelled words. If users fail to correct spelling errors, systems could check queries against go/see lists to determine if queries bear listed words or phrases. They would then enhance queries with words and phrases from go/see lists and start at the initial search tree to find matches. Section 12.1.5 discussed candidate types of terms and phrases for inclusions in go/see lists. Examples are one-word/two-word variations of concepts and noun/adjective pairs. If systems enhance queries with words and phrases from go/see lists, search trees start begin with the initial search tree and try to effect matches of the system-enhanced user query.

The alphabetical approach has been added to the search tree for one-word queries as the search type "of last resort." If all other subject searches fail to produce retrievals, systems should respond with the results of the alphabetical approach, i.e., a display of subject headings and *see* references in the alphabetical neighborhood of the user-entered term. If, however, the user query matched the criteria for an alphabetical search following the exact approach, the search type "of last resort" is the keyword-in-record search.

12.4.3 One-word Search Tree for Small Databases

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Expanding could be especially tiresome in the event that expanding produces few new titles that users have not retrieved in the results of previous subject searches. System designers could consolidate title-keyword and keyword-in-record searches into a single keyword search. The search tree for one-word queries would feature three subject searching approaches for small databases: (1) exact, (2) alphabetical, and (3) keyword-in-record searches. The latter should feature relevance feedback to find additional titles. Relevance feedback would involve an exact search for subject headings common to titles judged useful by system users.

12.4.4 Search Trees for Multi-word Queries

The redesigned search tree for multi-word queries features the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names (figure 12.3), (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type "of last resort." The redesigned tree is shown in figures 12.6A–C. The search tree for multi-word queries begins with system efforts to invoke the alphabetical approach (figure 12.6A).

If user queries fail to meet the criteria for the alphabetical approach, systems perform keyword-in-record searches for the individual words of user queries. If systems fail to produce retrievals for one or more query words, they present the words to users for spelling correction (figure 12.6C). Systems could assist users in correcting spelling errors using spelling-correction routines similar to such routines for word processing programs in which they suggest alternate spellings for misspelled words. If users fail to correct spelling errors, systems could check queries against go/see lists to determine if queries bear listed words or phrases (figure 12.6C). They then enhance queries with words and phrases from go/see lists and start at the initial search tree to find matches.

If the individual words in user queries produce retrievals in keyword-in-record searches, systems do not immediately show users the results (figure 12.6A). Instead, they pass queries onto other controlled vocabulary searches — the keyword-in-main-heading and keyword-in-subdivided-heading searches (figure 12.6A). If these searches produce too few or no retrievals, systems submit queries to free-text approaches (figure 12.6B).

Free-text approaches begin with the title-keyword search (figure 12.6B). If searchers want to find additional titles, systems can invoke relevance feedback in which they retrieve additional titles based on searches for subject headings common to titles users rated useful in title-keyword searches. If title-keyword searches fail to produce retrievals, systems can try keyword searches of subject heading fields and keyword-in-record searches. Following successful matches, systems can invoke relevance feedback using subject headings common to titles users rate useful to find additional titles. Of course, relevance feedback assumes that systems collect relevance assessments during the display of retrieved titles.

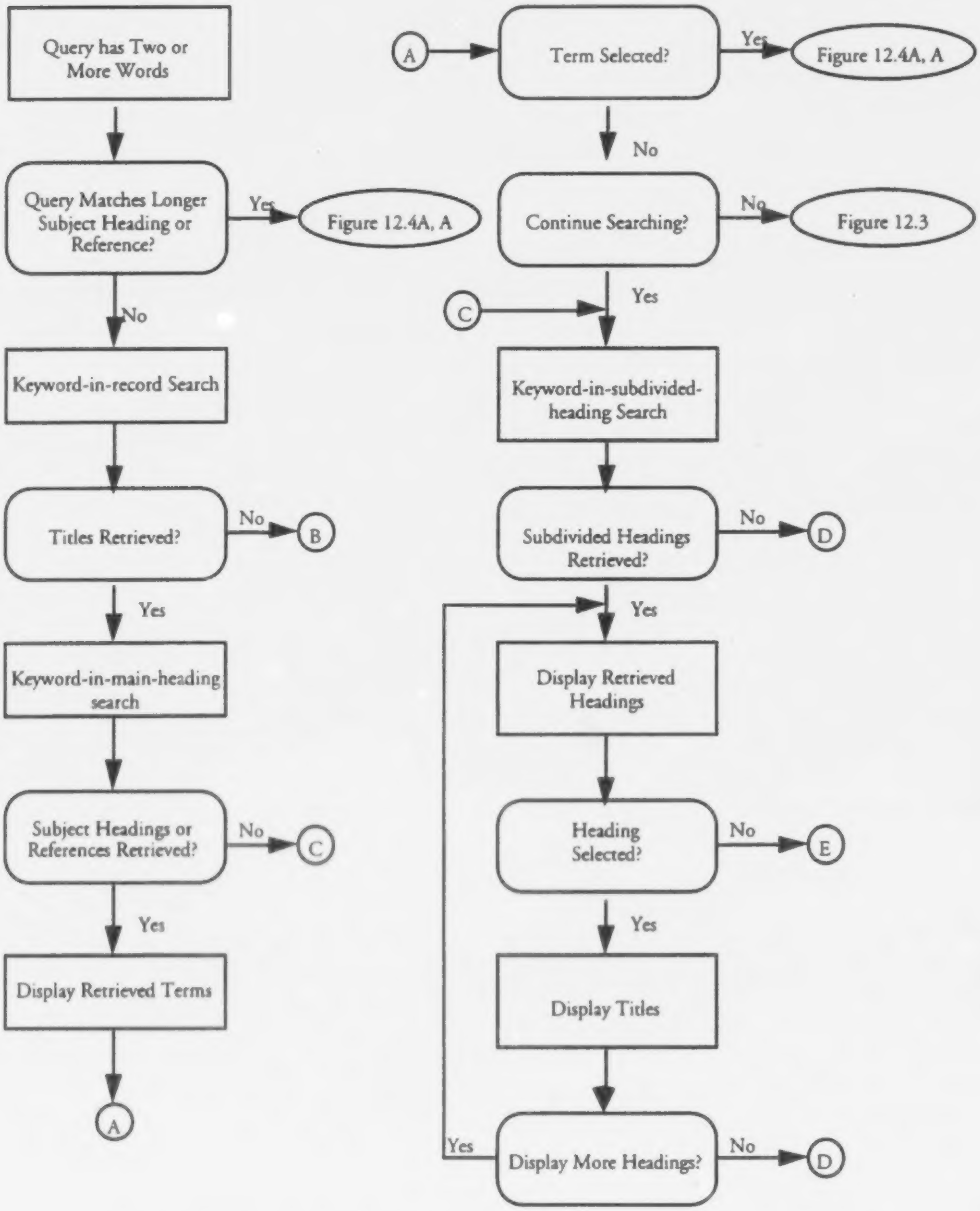

Query has Two or More Words
Query Matches Longer Subject Heading or Reference?
Yes
Figure 12.4A, A
No
Keyword-in-record Search
Titles Retrieved?
No
B
Yes
Keyword-in-main-heading search
Subject Headings or References Retrieved?
No
C
Yes
Display Retrieved Terms
A
A
Term Selected?
Yes
Figure 12.4A, A
No
Continue Searching?
No
Figure 12.3
Yes
C
Keyword-in-subdivided-heading Search
Subdivided Headings Retrieved?
No
D
Yes
Display Retrieved Headings
Heading Selected?
No
E
Yes
Display Titles
Display More Headings?
Yes
No
D

Figure 12.6A. Redesigned search tree for multi-word queries

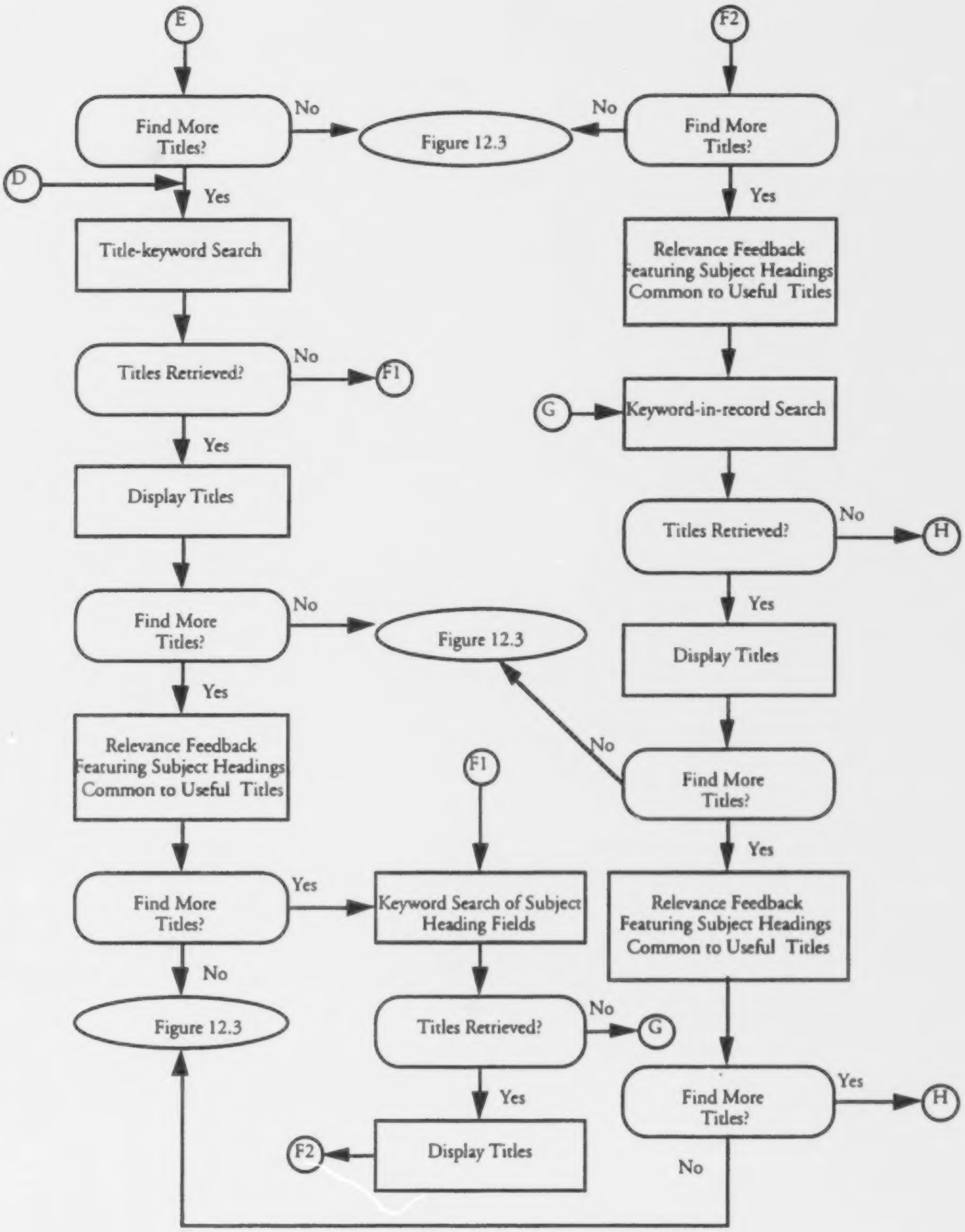

E
F2
Find More Titles?
No
Figure 12.3
No
Find More Titles?
D
Yes
Yes
Title-keyword Search
Relevance Feedback Featuring Subject Headings Common to Useful Titles
Titles Retrieved?
No
F1
G
Keyword-in-record Search
Yes
Display Titles
Titles Retrieved?
No
H
Yes
Find More Titles?
No
Figure 12.3
Display Titles
Yes
No
Relevance Feedback Featuring Subject Headings Common to Useful Titles
F1
Find More Titles?
Yes
Find More Titles?
Yes
Keyword Search of Subject Heading Fields
Relevance Feedback Featuring Subject Headings Common to Useful Titles
No
Figure 12.3
Titles Retrieved?
No
G
Yes
Find More Titles?
Yes
H
F2
Display Titles
No

Figure 12.6B. Redesigned search tree for multi-word queries (contd.)

BEST COPY AVAILABLE

Following the keyword-in-record search are two new approaches (figure 12.6C). One approach involves stemming of query words. The order of indexes that systems would search for stemmed query words are exact, keyword-in-main-heading, keyword-in-subdivided-heading, title-keyword, keyword in subject heading fields, and keyword-in-record. If systems are unsuccessful making matches of user queries, they should continue searching with the best-match approach.

The best-match approach would feature stemming, weighted-term probabilistic retrieval, and output ranking. Systems would perform searches on a best match, combinatorial basis. A match on fewer than all query-word stems may retrieve titles but titles that have all the query-word stems would be displayed first.

Query-word stems would be weighted so that frequently-occurring stems would get a low "importance" weight compared to the weight assigned to rare word stems. A new weighting system should be devised for library cataloging databases because subject content is represented in so few fields of bibliographic records, i.e., title and subject heading fields. Systems should increase the weight given to frequently-occurring words in subject headings and subdivisions because matches on such words are likely to retrieve subject headings describing users' topics of interest or a facet of their topics of interest.

The alphabetical approach has been added to the search tree for multi-word queries as the search type "of last resort." If all other subject searches fail to produce retrievals, systems should respond with the results of the alphabetical approach, i.e., a display of subject headings and *see* references in the alphabetical neighborhood of the user-entered term.

12.4.5 Multi-word Search Tree for Small Databases

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Expanding might be especially tiresome in the event that expanding produces few new titles that users have not retrieved in the results of previous subject searches.

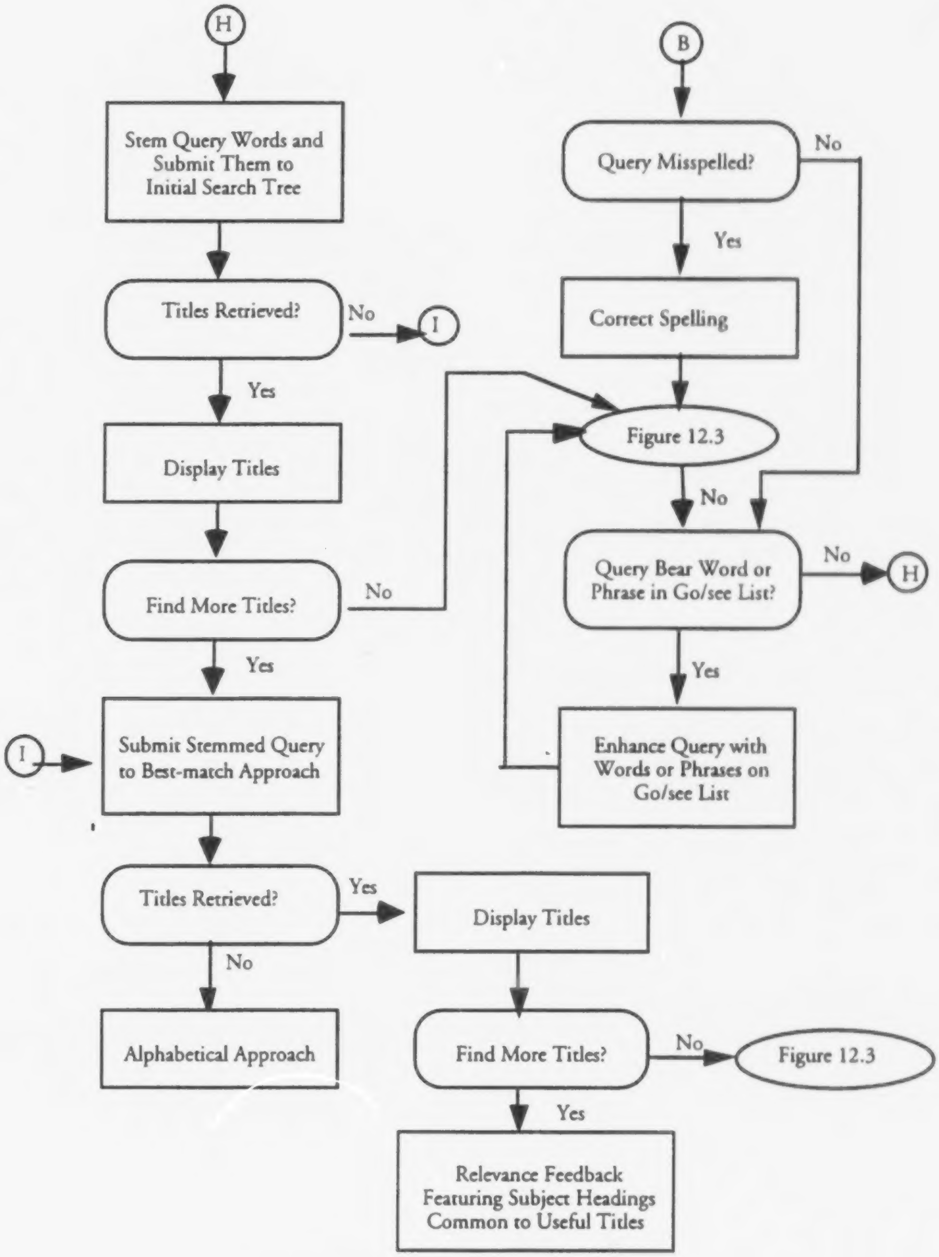

H
Stem Query Words and Submit Them to Initial Search Tree
Titles Retrieved?
No
I
Yes
Display Titles
Find More Titles?
No
Yes
I
Submit Stemmed Query to Best-match Approach
Titles Retrieved?
Yes
No
Alphabetical Approach
Display Titles
Find More Titles?
No
Figure 12.3
Yes
Relevance Feedback Featuring Subject Headings Common to Useful Titles
B
Query Misspelled?
No
Yes
Correct Spelling
Figure 12.3
No
Query Bear Word or Phrase in Go/see List?
No
H
Yes
Enhance Query with Words or Phrases on Go/see List

Figure 12.6C. Redesigned search tree for multi-word queries (contd.)

System designers could consolidate the results of the following searches: (1) keyword-in-main and keyword-in-subdivided-heading searches, (2) title-keyword, keyword in subject heading fields, and keyword-in-record searches, and (3) stemming and best-match approach. The search tree for multi-word queries would feature five subject searching approaches for small databases: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record, and (5) best match. Keyword-in-record and best-match searches would still feature relevance feedback to enable users to find additional titles on their topics of interest. Relevance feedback would involve an exact and keyword searches for subject headings common to titles judged useful by system users in keyword-in-record and best-match searches.

12.4.6 Search Trees for Subject Queries for Personal Names

The search tree for personal-name subject queries remains unchanged from the original conception (see Table 4.1). The search trees should be accompanied by pop-up windows that prompt users for the elements of their personal-name queries. This prompting enables systems to differentiate between name and topic elements of queries and choose to ignore certain elements in their attempts to retrieve titles. More research is needed to ensure that users enter the correct elements in various pop-up windows but this should not deter system designers from implementing the search trees and prompting mechanisms in operational systems.

The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching through the alphabetical approach. Personal name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

12.5 Chapter Summary

The analysis of online retrieval test data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. However, the search trees require enhancements to enable them to respond with useful retrievals to especially difficult user queries, that is, those for which the existing set of search trees produce no retrievals.

Chapter 12 focuses on the redesign of search trees for subject searching in online catalogs. Two problems — spelling and vocabulary problems — were greatly responsible for forcing a redesign of the search trees (sections 12.2 and 12.3). The analysis of spelling errors in user queries demonstrated that online catalogs should feature computer-assisted spelling correction routines — just like word processing programs — to assist end users in identifying and correcting their spelling errors. Solving vocabulary problems would require several enhancements to search tree configurations: (1) eliminating punctuation in queries including double posting hyphenated words in subject headings and title fields, (2) creating a go/see list to treat words and phrases as equivalent terms during retrieval, (3) automatic stemming of queries to effect retrievals in exact and the various keyword searches, and (4) best-match approach to retrieve titles for the most difficult user queries that traditional searches — exact, alphabetical, and keyword — cannot handle.

The only change to the initial search tree from the original design was returning users whose searches failed to produce retrievals to the question about personal names (see figure 12.3).

The search tree for the exact approach remained unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach (figures 12.4A and 12.4B).

The redesigned search tree for one-word queries contained a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type "of last resort." The search tree for one-word queries is given in figures 12.5A and 12.5B.

The redesigned search tree for multi-word queries featured the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type "of last resort." The redesigned tree is shown in figures 12.6A–C.

The search tree for personal-name subject queries remained unchanged from the original conception (Table 4.1). The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching through the alphabetical approach. Personal name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

Chapter 12 also contains recommendations for abridged search trees for subject searching in small databases (sections 12.4.3 and 12.4.5). Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Abridged search trees for one-word queries feature only three subject searching approaches: (1) exact, (2) alphabetical, and (3) keyword-in-record searches with relevance feedback. The search tree for multi-word queries would feature five subject searching approaches for small databases: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record with relevance feedback, and (5) best match with relevance feedback.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. Using subject headings for online *retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Walker, Stephen, and Richard M. Jones. 1987. *Improving subject retrieval in online catalogues; 1. Stemming, automatic spelling correction and cross-reference tables.* London: British Library. British Library Research Paper no. 24.

13 Highlights of Project Activities and Findings

13.1 Project Overview

This report describes the activities, analyses, findings, and recommendations of a research project that tests a new design for subject access to online catalogs. The foundation for the new design were findings from an empirical study of the subject terms users entered into online catalogs (Drabenstott and Vizine-Goetz 1994). The new design required that online catalogs have a wide range of subject searching capabilities and search trees to govern the system's selection of searching capabilities in response to user queries. Search trees hold much promise for assuming the burden of determining which subject searching approach is likely to produce useful information in response to user queries.

Project objectives were to:

1. Demonstrate subject searching in an experimental online catalog with a wide range of subject searching capabilities and with search trees to govern system selection of a particular capability in response to user queries.
2. Test the retrieval effectiveness of the experimental online catalog with search trees by comparing its performance with the performance of an experimental online catalog in which subject searching approaches are assigned at random.
3. Evaluate the demonstration and test results of retrieval effectiveness and disseminate the research findings through publications in the professional literature.

The research project was composed of twelve activities grouped into three phases (figure 1.1). Seven activities (1–6, 9) were required in the system-development phase. This phase featured the development of the ASTUTE experimental online catalog with a wide range of subject searching functionality and search trees to govern the system's selection of a subject searching approach in response to user queries. The databases of the experimental online catalog contained bibliographic records in computer science and technology from Mardigian Library of the University of Michigan-Dearborn (UM-D) or bibliographic records in American history from Lilly Library of Earlham College, and subject authority records from the Library of Congress' subject authority file. Three activities (7–8, 10) were required in the system testing phase. Data collection procedures and instruments were developed, pretested, and revised. Instruments were administered in online retrieval experiments with library patrons and staff at Mardigian and Lilly Libraries of UM-D and Earlham College to answer the following three research questions:

1. Do search trees improve the search performance of subject searchers at the online catalog?
2. Do subject searchers prefer an online catalog that controls system responses and searching approaches to an online catalog without such controls?
3. What are the characteristics of queries that cannot be answered using the ASTUTE experimental online catalog?

Two activities (11–12) made up the evaluation phase of the project. Collected data were described, analyzed, interpreted, and synthesized in a final report to the Department of Education. The ASTUTE Project Team also plans to disseminate findings in journals and conference proceedings to the library and information science communities.

13.2 Development of the ASTUTE Experimental Online Catalog

13.2.1 Computer Equipment

ASTUTE (A Search Tree Underlying the Experiment) was programmed on a stand-alone Gateway 2000 486, 33 MHz, IBM-compatible microcomputer, with 8 megabytes of RAM and a VGA color monitor (section 3.2). The operating system was MS-DOS version 5.0. A dot-matrix printer and a mouse were attached to the microcomputer for use by ASTUTE project staff during development work and during online retrieval tests with end users and librarians.

When ASTUTE was installed in the libraries of UM-D and Earlham College, the system did not require a network connection. It resided entirely on the Gateway microcomputer. Participating library staff monitored the system, performed daily system backups, and periodically used the microcomputer equipment and network connections in their institutions to transmit transaction log files of end-user search activity to the ASTUTE project team in Ann Arbor.

13.2.2 Database Development

The databases of the ASTUTE experimental online catalog were created from two data sources: (1) machine-readable cataloging (MARC) records for bibliographic data from the two participating libraries (section 3.3), and (2) USMARC records for subject authority data from the compact disc-based product CD/MARC Subjects distributed by the Library of Congress (section 3.4). Bibliographic records from the University of Michigan-Dearborn and Earlham College were limited to computer science and technology, and American history, respectively. Subject authority records were limited to subject headings assigned to bibliographic records.

The team made a few changes to bibliographic data: (1) deleting diacritics in Earlham bibliographic records (section 3.3.3.1), and (2) correcting erroneous subject headings (section 3.3.3.2). The few enhancements the team made to bibliographic data were intended to enhance subject access: (1) identifying and coding form subdivisions to increase the number of broad categories in exact search from three to four (section 3.3.4.1), and (2) converting control-field codes to English-language equivalent terms to increase the searchable dictionary in keyword-in-record searches (section 3.3.4.2). The team deleted diacritics in subject authority records (section 3.4.2) and enhanced records with narrower terms (section 3.4.3).

13.2.3 Search Trees and System Functionality

The ASTUTE experimental online catalog was actually composed of two online catalogs: (1) the Blue System, in which search trees governed the system's selection of a subject searching capability, and (2) the Pinstripe System, in which the system selected a subject searching capability randomly. These systems were purposely designed to be very much alike to focus the attention of library patrons and staff on the retrieval of useful information in response to their queries. The Blue and Pinstripe Systems had virtually the same interfaces, and they accessed the same bibliographic and authority databases. Except for the Blue System's enhancement with the search trees, the two systems and their capabilities were very much the same.

The Blue and Pinstripe Test Systems used different criteria to choose subject searching approaches in response to user queries. Search trees governed the Blue System's selection of subject searching approaches (sections 4.2–4.5). Search trees were defined as a set of paths with branches or choices that enabled systems to carry out the most sensible search approach at each stage of the search (Mitev, Venner, and Walker 1985). The search trees incorporated into the Blue System emanated from the findings of an empirical study of user queries (Drabenstott and Vizine-Goetz 1994). The Pinstripe System's selection of a subject searching approaches was governed by a random selection algorithm (section 5.10).

Search trees preferred two-step subject searching approaches that enlisted the catalog's controlled vocabulary, i.e., exact, alphabetical, and keyword-in-heading search approaches, before free-text approaches, i.e., title-keyword, keyword searches of subject heading fields, and keyword-in-record searches. If the former approaches supplied users with appropriate controlled vocabulary terms for expressing their topics of interest, they provided users with several opportunities to refine their searches using related terms, subdivided forms of the matched or selecting heading, or controlled vocabulary terms in close alphabetical proximity to the subject queries users entered into online catalogs.

The Blue System featured four controlled vocabulary approaches: (1) exact (subjects only), (2) alphabetical (subjects and personal names), (3) keyword-in-subdivided-heading (subjects only), and (4) keyword-in-main-heading (subjects only) searches. It featured three free-text approaches: (1) title-keyword (subjects only), (2) keyword in subject heading fields (subjects only), and (3) keyword-in-record (subjects and personal names) searches.

The Pinstripe System's selection of a subject searching approaches was governed by a random selection algorithm (section 5.10). This algorithm chose between three different subject searching searches: (1) alphabetical (subjects only), (2) keyword-in-subdivided-heading (subjects only), and (3) keyword-in-record (subjects and personal names).

Both Blue and Pinstripe Systems prompted users for the last name, first name, and topic elements of their subject queries for personal names (section 5.9.2). The Pinstripe System always responded to such queries with the keyword-in-record search. The Blue System preferred the keyword-in-record search for personal-name queries. If it was unsuccessful retrieving titles, it discarded the topic element and submitted personal-name queries to the alphabetical approach.

13.3 Methods Used to Test the New Subject Access Design

The methods used in this project to compare the performance of the Blue and Pinstripe Systems were based on the Comparison Search Experiment devised by Siegel et al. (1983) in their evaluation of two prototype online catalogs at the National Library of Medicine (NLM) (section 6.2). Comparative approaches to system evaluation like Siegel's and several modeled after Siegel's experiment have used human intermediaries to collect information about searchers and their searching experiences. Such approaches are expensive and time-consuming for several reasons (section 6.3). For example, interviewers must be present to monitor online searches, ask users questions, and record their answers. The project director used a comparative approach to system evaluation in this project that did not use interviewers or human intermediaries at all. The experimental online catalog performed and/or recorded all activity connected with data collection — searcher recruitment, questionnaire administration, search logging, and searcher relevance assessments.

In the Comparison Search Experiment, the Blue and Pinstripe Systems encouraged library patrons and staff to perform online subject searches in both systems (section 6.4.1). In introductory screens, the systems informed users that they were participating in an experiment and allowed them to graciously exit the system if they did not want to take part. Prior to conducting subject searches, the systems asked participants to respond to three pre-search questions that collected demographic information on participants (section 6.4.2). During their searches, the systems recorded user activity and user relevance assessments for displayed bibliographic records to a time-stamped transaction log (section 6.6 and Appendix F). The systems asked participants to answer eleven questions in a post-search questionnaire that compared the performance of the Blue and Pinstripe Test Systems (section 6.4.4 and Appendix C). Thus, the retrieval tests yielded quantitative comparative data about the systems' capabilities and effectiveness, and about users' system preference. After completing the post-search questionnaire, library staff were asked additional questions that prompted them to make comparisons between the two systems with respect to the systems' ease of use, subject searching capabilities, and performance (section 6.5 and Appendix E). These qualitative data were combined with quantitative data from patron searches to increase our understanding of system performance.

13.4 Comparison Search Experiment Participants

Before the researchers could submit Comparison Search administrations to the data analysis, they had to manually review individual search administrations to determine whether patron search topics were in scope and whether patron responses to pre- and post-search questionnaires were valid.

The three installations of the experimental online catalog lasted for five- to fifteen-week periods at the two participating libraries. Data collected in the first of two data-collection periods at UM-D were discarded because of a system error that corrupted data for over two-fifths of the search administrations (section 7.4). The system error was introduced by the unexpected user action of repeatedly turning the microcomputer on and off. The programming team made changes to the experimental online catalog to ensure that the data would not become corrupted in later system installations.

In the second, five-week data-collection period at UM-D, the experimental online catalog administered a total of 826 search experiments (section 7.5). In the fifteen-week data-collection period at Earlham, the experimental online catalog administered a total of 238 search experiments (section 7.6).

The manual review of searches revealed that about 50% of search administrations at UM-D and Earlham were usable (sections 7.5.2 and 7.6.2); however, the manual review of post-search responses revealed that large percentages (58% at UM-D, 51% at Earlham) of search experiments were less-than-full administrations. That is, users completed a search in one system, both systems, and/or failed to answer the post-search questionnaire. Since monitors were not present to encourage patrons to complete the entire Comparison Search Experiment, the researchers expected patrons to leave before they finished the full administration of the Comparison Search Experiment. The manual review process was essential because it identified full and less-than-full search administrations which were, on occasion, handled differently in the statistical analysis of search results. Only searches with usable queries were submitted to statistical analyses and failure analyses (chapters 8–12).

Unusable administrations were divided into categories for queries that were out of scope, for play, meaningless input, known-item searches, and other unusable searches (e.g., sex terms, expletives) (sections 7.5.3 and 7.6.3). Much of sections 7.5 and 7.6 was given to discussions of unusable search administrations — systems searched, titles retrieved, relevance assessments, etc. Discouraged with the preponderance of positive relevance assessments users gave to displayed titles in unusable search administrations, the researchers made suggestions for redesigning the procedures for collecting relevance assessments in future experimental and operational systems (section 7.7). They also made suggestions for redesigning system-administered post-search questionnaires to minimize the collection of unreliable data in future experimental systems (section 7.8).

13.5 Characteristics of System Users and Searches

13.5.1 System Users

The majority of users at both UM-D and Earlham were searching ASTUTE for the first time (section 8.2). The majority were also daily users of computer systems. Large percentages of UM-D respondents were computer science or engineering majors (29.1% and 46.1%, respectively). Thus, the majority of UM-D respondents were performing searches for topics in their major field of study. This was not true of respondents at Earlham College where ASTUTE's database focused on American history. Only one-quarter of Earlham respondents were history majors; consequently, the majority of Earlham respondents were not performing searches for topics in their major field of study.

Searchers did not spend much time using the ASTUTE experimental online catalog (section 8.3). They spent less than a minute completing the pre-search questionnaire. Time spent searching was quite low for search administrations bearing suspect post-search questionnaires — a little more than three minutes at UM-D and two minutes at Earlham. Search administrations were the lengthiest for complete search administrations — over four minutes at both UM-D and Earlham. Completing post-search questionnaires took less than a minute in search administrations with suspect questionnaire responses and took a minute and a half in full search administrations.

13.5.2 Searches

ASTUTE was designed to handle one query per search experiment administration (section 8.4). Thus, the average number of queries per search (1.2) was much lower than the average number in operational online catalogs. A statistical test demonstrated that the length of queries had a significant effect on the ability of the experimental online catalog to make matches of its subject vocabulary (Table 8.5). Long user queries (three or more words) were less likely to make matches than short queries (two or fewer words).

13.5.3 Matches of User Queries

The first comparison of system performance was given in an analysis of system responses that designated matches of user queries (section 8.5.1). Matches gave users opportunities to pursue the results of subject searching approaches. Generally, the Blue System was likely to provide more opportunities for guiding users to useful information on their topics of interest than the Pinstripe System provided (Table 8.6).

Providing users the opportunity to display titles did not guarantee that users would, especially in the case of controlled vocabulary approaches, choose a seemingly relevant subject heading from a list that would take them one step closer to retrieving titles connected with the subject heading. Furthermore, it did not guarantee that users would browse more than the first title displayed in free-text approaches. Users seldom displayed all the titles they retrieved, in fact, the statistics in Table 8.8 showed that they displayed about a fifth of the titles they retrieved. At UM-D, the Pinstripe System retrieved $2^1/_2$ times as many titles as the Blue System. High-posted searches in the Pinstripe System were likely to come from the keyword-in-record search; this search made retrievals based on matches in all subject-rich fields. Pinstripe-system users at UM-D displayed only about an eighth of the titles they retrieved.

13.5.4 Resolving Conflicting Relevance Assessments

At UM-D, Blue- and Pinstripe-system users displayed a total of 470 titles more than once. Of this total, 82.8% of titles were displayed twice in across systems, 11.5% were displayed three times across systems, and 5.7% were displayed more than twice across systems. At Earlham, Blue- and Pinstripe-system users displayed a total of 176 titles more than once. Of this total at Earlham, 80.1% of titles were displayed twice across systems, 12.5% were displayed three times across systems, and 7.4% were displayed more than three times across systems. The researchers reviewed relevance assessments for the same titles and noticed that the more users displayed the same titles, the more likely they were to give positive relevance assessments to titles, i.e., "very useful" or "possibly useful" ratings. After much discussion among team members and consultants, we decided to use the last relevant assessment that searchers gave to titles displayed more than once and to disregard other relevance assessments for the same title. The resolution of such conflicts did not have a significant effect on sets of rated titles (Table 8.10). Resolution was necessary, however, to submit the data to statistical procedures to measure precision and estimated recall.

13.6 Precision and Estimated Recall in the Blue and Pinstripe Systems

13.6.1 Precision Scores

At UM-D, precision scores averaged 64% in the Blue System and 58% in the Pinstripe System (section 9.2.1). In full search administrations at UM-D (i.e., administrations in which users conducted searches in both Blue and Pinstripe Systems and completed a post-search questionnaire), precision scores in the Blue System were significantly higher than precision scores in the Pinstripe System (figure 9.1. and Table 9.1). The result was exactly the opposite at Earlham. That is, precision scores in the Pinstripe System were significantly higher than precision scores in the Blue System (figure 9.2. and Table 9.2).

We extended the analysis of precision scores to comparisons for controlled vocabulary and free-text searches because the Blue System's search trees favored controlled vocabulary over free-text searches in response to user queries (section 9.2.2). We expected the precision of titles retrieved through controlled vocabulary searches to exceed the precision of titles retrieved through free-text searches. Such results would support the design of search trees in the experimental online catalog.

In UM-D's Blue and Pinstripe Systems, precision scores for controlled vocabulary searches exceeded such scores for free-text searches (figures 9.3 and 9.4). Furthermore, the differences between scores were significant (Tables 9.3 and 9.4). Especially large differences between precision scores were registered for full search administrations. In Earlham's Blue System, precision scores for controlled vocabulary approaches exceeded such scores for free-text approaches (figure 9.5). Again, the difference between scores was especially large for full administrations. There was an insufficient number of searches for statistical comparisons of precision in Earlham's Pinstripe System; however, overall precision scores for controlled vocabulary approaches were significantly higher than such scores for free-text approaches (Table 9.6). These findings about precision and search approaches support the design of search trees for searching bibliographic databases that favor controlled vocabulary over free-text approaches.

13.6.2 Recall Scores

Overall, recall scores in UM-D's experimental online catalog were about the same (section 9.3). Recall in the Blue System was 53.4% and recall in the Pinstripe System was 55.4% (figure 9.6). Recall scores in the two search administrations flip flopped. That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Type of search administration results were the same at Earlham — recall scores in the two search administrations flip flopped (figure 9.7). That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Overall, recall in Earlham's Pinstripe System (61.8%) was somewhat higher than recall in the Blue System (50.8%). No statistical tests were performed on recall scores because the design of the comparison search experiment violated the t-test's principle of independence.

13.7 Post-search Questionnaire Responses Regarding System Performance

End-user responses to the post-search questionnaire were available for full search administrations (section 9.4). Generally, responses reflected statistical results of this type of search administration. That is, users rated the performance of the Blue System higher than the Pinstripe System in terms of retrieving greater numbers of useful titles (figure 9.8), and search satisfaction (figure 9.9), and they preferred the Blue System over the Pinstripe System (figure 9.10).

The post-search questionnaire probed other factors — ease of system use, system efficiency in retrieving useful titles, system helpfulness in giving users new subject searching ideas — in the comparison of Blue and Pinstripe Systems. The majority of end users rated the Blue System or both systems highly in terms of the ease of giving instructions to the systems (figure 9.11), ease of getting instructions from the system (figure 9.12), system clarity (figure 9.13), and system efficiency in guiding them to useful information (figure 9.14).

At Earlham, users were a little less enthusiastic about the Blue System's helpfulness in giving new ideas for subject searching (figure 9.15). Since the design of the Blue System was intended to give users new subject searching ideas through exact-approach browsing and expanding opportunities, we were somewhat concerned about this response. The results of a full-scale failure analysis of experimental catalog searches revealed the reasons why users answered the question in this way and provided many areas for improving this approach (chapters 10–12).

13.8 Failure Analysis of User Searches

13.8.1 Characterizing Successful and Unsuccessful Searches

This study featured a failure analysis of end-user searches to determine why the experimental ASTUTE online catalog was successful or unsuccessful in retrieving useful titles. The project director, an expert searcher of information retrieval systems, repeated end-user searches that were successful and unsuccessful in retrieving useful titles to identify the reasons why they succeeded or failed.

The first step in the failure analysis of searches performed at the University of Michigan-Dearborn and Earlham College was to characterize searches that users performed in the Blue and/or Pinstripe Systems as successful or unsuccessful:

- Successful search — users rated one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful search — users did not rate one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful pairs of searches — users did not rate titles "very useful" in searches performed in the Blue *and* Pinstripe Systems for the *same* topic.

At both UM-D and Earlham (sections 10.2 and 10.3), about 44% of searches conducted in the Blue System and 37% to 43% of searches conducted in the Pinstripe System were successful. About 19% to 25% of searches conducted in the Blue System or Pinstripe System were unsuccessful. About a third of searches conducted in the Blue *and* Pinstripe Systems for the *same* topic were unsuccessful.

13.8.2 Reasons for Failure

The failure analysis of these searches demonstrated that the following problems adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failure, (5) large retrievals, (6) user query vocabulary, (7) search approach failures, (8) navigational problems, and (9) spelling problems. Sections 10.2 and 10.3 described the failure analysis of searches conducted by users at UM-D and Earlham, respectively.

13.8.3 Reclassifying Successful Searches

Explanations and examples of searches placed in categories for perseverance, query specificity, and conflicting relevance assessments provided evidence that these searches were really successful searches in terms of leading users to titles on their topics of interest. With regard to user perseverance, the systems gave users ample opportunities for browsing terminology or displaying titles on their topics of interest but users terminated their searches. In searches marred by query specificity, users gave negative relevance assessments to the titles that seemed relevant to their queries. Their actions led us to conclude that the systems were successful in terms of retrieving titles on the topics users expressed in their queries but they were not successful in terms of retrieving titles that fulfilled the specifications of queries users had in their minds but did not make explicit in their queries. In the failure analysis, we also considered searches bearing conflicting relevance assessments as successful searches because users had given the titles positive assessments in the course of their searches.

The reclassification of searches placed into categories for perseverance, query specificity, and conflicting relevance assessments from unsuccessful to successful searches resulted in success rate of about 72% for Blue-system searches at UM-D and Earlham and a success rate of between 49% and 59% for Pinstripe-system searches at UM-D and Earlham (Tables 10.17 and 10.31). Clearly, the Blue System performed better than the Pinstripe System in terms of retrieving and displaying titles on the topics users made explicit in their queries (sections 10.2.3 and 10.3.3).

13.8.4 Associating Specific Failures with Subject Searching Approaches

Success in the Blue System, however, was marred by a large percentage of searches in which users did not persevere in terms of browsing subject terminology and displaying retrieved titles (sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1). Almost ten percentage points separated the results of Blue- and Pinstripe-system searches for searches marred by perseverance (Tables 10.17 and 10.31). Many Blue-system searches affected by perseverance featured controlled vocabulary search approaches. Navigational problems were also linked to user perseverance and controlled vocabulary searches (sections 10.2.1.8, 10.2.2.8, and 10.3.2.8). Of the several suggestions made to improve controlled vocabulary searches, the important ones were:

- Reducing the response time for searching selected terms and creating exact-search option menus.
- Reserving exact-search option menus for summaries of large retrievals and seeking other methods for summaries of one or a few retrievals.

Success in the Pinstripe System was negatively affected by the system's random selection algorithm for subject searching approaches. Between ten and twenty percentage points separated the results of Blue- and Pinstripe-system searches for search approach failures (Tables 10.17 and 10.31). The keyword-in-subdivided-heading search was usually to blame for search approach failures; however, the alphabetical approach had a large share of the failures. Subject searching approach failures were rare in the Blue System, accounting for only six failures in over 450 searches. The Blue System was guided in its selection of a subject searching approach by the search trees. Thus, the data and analyses in this project provided evidence that search trees improved the search performance of subject searchers at the online catalog in terms of guiding them to subject headings and titles that described the topics they made explicit in their queries.

13.9 Failure Analysis of Subject Searching Approaches

13.9.1 Controlled Vocabulary Approaches

Controlled vocabulary searches were especially plentiful in the Blue System and accounted for almost two-thirds of the searches conducted in this system (section 11.2). About half (52.5%) of controlled vocabulary searches were successful in retrieving at least one title rated "very useful" by system users. Many (54.5%) of these successful searches were searches characterized by no browsing. That is, users displayed retrieved titles and terminated their searches without browsing available options for broader or narrower terms, subdivided terms, notes, or expanding the search. Only 14.4% of successful controlled vocabulary searches featured extensive browsing (Table 11.5).

Frequent reasons for the failure of controlled vocabulary searches in the Blue System were perseverance and navigation problems (sections 11.2.4 and 11.2.5). Some of the searches in which perseverance was a problem could have been navigation problems. Based on the analysis of searches that failed due to perseverance and navigation problems, the following suggestions were made for improving the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.
- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

The keyword-in-subdivided-heading search was a controlled vocabulary search but it was treated separately from other controlled vocabulary searches because it was independent of the exact approach (section 11.3). Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing (section 11.3.3). These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing.

Unsuccessful keyword-in-subdivided-heading searches outnumbered successful keyword-in-subdivided-heading searches in the Pinstripe System (section 11.3.4). Frequent reasons for failure were perseverance, database failure, and searches that retrieved no subject headings.

13.9.2 Free-text Approaches

Title-keyword searches were unique to the Blue System (section 11.4). Successful title-keyword searches outnumbered unsuccessful title keyword searches. Almost 70% of title keyword searches in UM-D's Blue System were successful; over 85% of title keyword searches in Earlham's Blue System were successful. On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. When title-keyword searches retrieve few titles, users could find additional material on their topics of interest from the system's retrieval of titles bearing these subject headings.

Both Blue and Pinstripe Systems featured keyword-in-record searches but the majority (95.3%) of keyword-in-record searches were conducted in the Pinstripe System (section 11.6). About half of keyword-in-record searches were successful. Keyword-in-record searches were low posted in the Blue System where they averaged 2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Large retrievals were the major problem with keyword-in-record searches in the Pinstripe System. Often users walked away without rating displayed titles or terminated their searches after rating the first of many retrieved titles.

13.9.3 Subject Searching for Personal Names

No subject queries for personal names that users entered into the experimental online catalog at UM-D were usable names (section 11.7.2.1). Most were out-of-scope or playing around. Users also entered subject queries for personal names that were deemed unusable into the catalog at Earlham (section 11.7.2.2); however, a total of 27 usable queries were entered. All but four of these queries contained topical elements. Blue-system search trees for handling personal-name queries were especially effective in retrieving useful titles for user queries (section 11.7.2.4).

13.9.4 Spelling and Vocabulary Problems

The analysis of spelling errors in user queries demonstrated that online catalogs should feature computer-assisted spelling correction routines — just like word processing programs — to assist end users in identifying and correcting their spelling errors (section 12.2). Solving vocabulary problems (section 12.3) would require several enhancements to search tree configurations: (1) eliminating punctuation in queries including double posting hyphenated words in subject headings and title fields, (2) creating a go/see list to treat words and phrases as equivalent terms during retrieval, (3) automatic stemming of queries to effect retrievals in exact and the various keyword searches, and (4) best-match approach to retrieve titles for the most difficult user queries that traditional searches — exact, alphabetical, and keyword — cannot handle.

13.10 Redesigning Search Trees for Responsive Subject Searching

13.10.1 Redesigning Search Trees

Statistical and failure analyses resulted in many recommendations to improve the operation of specific subject searching approaches. They also resulted in recommendations to improve the design of search trees.

The only change to the initial search tree from the original design is to return users whose searches failed to produce retrievals to the question about personal names (see figure 12.3).

The search tree for the exact approach remains unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach (figures 12.4A and 12.4B).

The redesigned search tree for one-word queries contains a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type "of last resort." The search tree for one-word queries is given in figures 12.5A and 12.5B.

The redesigned search tree for multi-word queries features the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type "of last resort." The redesigned tree is shown in figures 12.6A–C.

The search tree for personal-name subject queries remains unchanged from the original conception (Table 4.1). The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching through the alphabetical approach. Personal-name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

13.10.2 Abridged Search Trees

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Thus, we recommend abridged search trees that do not feature as many subject searches and expanding options (sections 12.5.3 and 12.5.5). Abridged search trees for one-word queries feature only three subject searching approaches: (1) exact, (2) alphabetical, and (3) keyword-in-record searches with relevance feedback. The search tree for multi-word queries features five subject searching approaches: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record with relevance feedback, and (5) best-match with relevance feedback.

13.11 A New Design for Subject Access to Online Catalogs

This project tested a new design for subject access to online catalogs. It compared the performance of an experimental online catalog that was governed in its selection of subject searching approaches by search trees (Blue System) with the performance of an experimental online catalog that chose subject searching approaches randomly (Pinstripe System). The random selection of subject searching approaches was intended to reflect how end users select subject searching approaches when systems provide them with multiple choices.

The Blue System was more likely than the Pinstripe System to provide opportunities for guiding users to useful information on their topics of interest than the Pinstripe System provided. In full administrations of the Comparison Search Experiment at UM-D, precision of Blue-system searches was significantly greater than precision of Pinstripe-system searches. In both Blue and Pinstripe Systems at UM-D and Earlham, controlled vocabulary searches fared significantly better than free-text searches. About three-quarters of Blue-system searches were successful in retrieving at least one title that users judged "very useful." This definition of successful searches included sizable numbers of Blue-system searches that were marred by user perseverance and navigation problems. Recommendations were made to improve the design of the exact approach and minimize problems connected with perseverance and navigation.

End users rated the performance of the Blue System much higher than the Pinstripe System in terms of retrieving greater numbers of useful titles and making them satisfied with their search, and they preferred the Blue System over the Pinstripe System.

Overall, between 50% and 60% of Pinstripe-system searches were successful. Success in the Pinstripe System was negatively affected by the system's random selection algorithm for subject searching approaches. Between ten and twenty percentage points separated the results of Blue- and Pinstripe-system searches for search approach failures. The keyword-in-subdivided-heading search was usually to blame for search approach failures; however, the alphabetical approach had large shares of the failures. Subject searching approach failures were rare in the Blue System, accounting for only 1.3% of Blue-system searches. Search-approach failures accounted for the failure of 14.3% of Pinstripe-system searches. Pinstripe-system searches were also negatively affected by large retrievals — a problem that seldom occurred in Blue-system searches.

The analysis of Comparison Search Experiment data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. This project provided recommendations for improving specific subject searching approaches to increase their efficiency, increase user perseverance, and encourage browsing. It also suggested enhancements to the search trees to enable them to respond to the wide variety of user queries for subjects.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential.* San Diego: Academic Press.

Mitev, Nathalie, Gillian Venner, and Stephen Walker. 1985. *Designing an online public access catalog.* London: British Library. Library and Information Research Report 39.

Siegel, Elliott R. et al. 1983. "Research strategy and methods used to conduct a comparative evaluation of two prototype online catalog systems." In *National Online Meeting proceedings —1983*, 1983 April 12–14, New York, compiled by Martha E. Williams and Thomas H. Hogan, 503–11. Medford, N.J.: Learned Information.

Appendix A. Form Subdivisions

Topical subdivisions assigned to subfield code $1 for form subdivisions in bibliographic records of the University of Michigan-Dearborn:

Abbreviations
Abstracts
Amateurs' manuals
Anecdotes, facetiae, satire, etc.
Atlases
Bibliography
Bio-bibliography
Biography
Book lists
Case studies
Catalogs
Charts, diagrams, etc.
Chronology
Congresses
Diagrams
Dictionaries
Directories
Drawings
Early works to 1800
Encyclopedias
Examinations, questions, etc.
Exhibitions
Film catalogs
Films
Formulae, receipts, prescriptions
Guide-books
Handbooks, manuals, etc.
Illustrations
Indexes
Job descriptions
Juvenile literature
Laboratory manuals
Maps
Miscellanea
Observers' manuals
Outlines, syllabi, etc.
Periodicals
Photograph collections
Pictorial works
Popular works
Portraits
Problems, exercises, etc.
Programmed instruction
Registers
Study guides
Tables
Textbooks for foreign speakers
Union lists

Topical subdivisions assigned to subfield code $1 for form subdivisions in bibliographic records of Earlham College:

Abstracts
Addresses, sermons, etc.
Anecdotes
Archival resources
Audio-visual aids
Bibliography
Bio-bibliography
Biography
Caricatures and cartoons
Cartoons, satire, etc.
Case studies
Cases
Catalogs
Chronology

Congresses
Controversial literature
Correspondence
Curricula
Designs and plans
Diaries
Dictionaries
Dictionaries, indexes, etc.
Directories
Early accounts to 1600
Early works to 1800
Encyclopedias
Exhibitions
Facsimiles
Festschriften
Fiction
Field work
Gazetteers
Glossaries
Glossaries, vocabularies, etc.
Guide-books
Handbooks, manuals, etc.
Humor, caricatures, etc.
Indexes
Interpretive programs
Interviews
Juvenile literature
Library resources
Lists
Manuscripts
Maps
Maps, Early American
Miscellanea
Outlines, syllabi, etc.
Pamphlets
Passenger lists
Periodicals
Photograph collections
Pictorial works
Pictures, illustrations, etc.
Popular works
Portraits
Portraits, caricatures, etc.
Registers
Registers of births, etc.
Reviews
Sermons
Sources
Statistics
Statistics, Vital
Tables
Textbooks
Texts
Vocabularies, etc.
Voting registers

Appendix B.
Control-field Elements

**Table B1. Other Contents: Illustration Code
(USMARC Tag 008, Positions 18-21)**

Code	English-language equivalent term
a	Illustrations
b	Maps
c	Portraits
d	Charts
e	Plans
f	Plates
g	Music
h	Facsimiles
i	Coats of arms
j	Genealogical tables
k	Forms
m	Phonodisc, phonowire, etc.
o	Photographs
p	Illuminations

**Table B2. Book Details: Nature of Contents Code
(USMARC Tag 008, Positions 24-27)**

Code	English-language equivalent term
t	Technical reports
b	Bibliographies
k	Discographies
q	Filmographies
c	Catalogs
i	Indexes

a	Abstracts/summaries
d	Dictionaries
e	Encyclopedias
r	Directories
s	Statistics
f	Handbooks
p	Programmed texts
w	Law reports and digests
g	Legal articles
o	Reviews
v	Legal cases and case notes
n	Literature surveys

Table B3. Book Details: Festschrift Code
(USMARC Tag 008, Position 30)

Code	English-language equivalent term
1	Festschrift

Table B4. Book Details: Biography Code
(USMARC Tag 008, Position 34)

Code	English-language equivalent term
1	Contains biographical information

Appendix C.
End-user Questionnaire

The pre-search questionnaire consists of questions 1–3. The post-search questionnaire consists of questions 4–14. Preceding questions 1, 4, 7, and 13 are comments that introduce several questions.

Please begin by answering three questions about yourself:

1. How many times have you tested this system?
   ( ) Never before
   ( ) One time
   ( ) Two times
   ( ) Three or more times
2. How often do you use other computer systems?
   ( ) Daily
   ( ) About every other day
   ( ) Weekly
   ( ) Monthly
   ( ) About four times a year
   ( ) About once a year
   ( ) Never
3. What is your major field of study?
   ( ) Computer science
   ( ) Engineering
   ( ) Business and management
   ( ) Education
   ( ) History
   ( ) Mathematics
   ( ) Other humanities fields

( ) Other sciences fields
( ) Other social sciences fields
( ) None of the above

Please answer some general questions about the titles the Blue and Pinstripe Test Systems found for the subjects you typed in.

4. Comparing the Blue Test System with the Pinstripe Test System, did you find:
( ) A greater amount of useful titles in the Blue Test System
( ) About the same amount of useful titles in both systems
( ) A lesser amount of useful titles in the Blue Test System
( ) No useful titles in either system

5. In relation to what you were looking for, would you say your search was more satisfactory in:
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

6. The next time you need to conduct a computer search, will you want to use:
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

Please answer the next six specific questions about using the Blue and Pinstripe Test Systems.

7. Which system do you find easier to give instructions to, for example, to type in subjects, continue searches, make selections from lists?
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

8. Which system gives you more useful instructions in response to the subjects you type in?
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

9. Which system tells you what to do more clearly?
( ) The Blue Test System

( ) The Pinstripe Test Systems
( ) No difference

10. Which system is more helpful for giving you new ideas for searching for subjects?
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

11. Which system is more efficient in guiding you to useful information?
( ) The Blue Test System
( ) The Pinstripe Test Systems
( ) No difference

12. On a scale of 1 to 5 in which 1 means "not at all important" and 5 means "very important," rate the following features as to their importance in helping you search library computer systems like ASTUTE:
( ) Ease of giving instructions to the system
( ) Usefulness of the system's instructions
( ) New ideas that the system gives you for searching for subjects
( ) The system's clarity in telling you what to do
( ) The system's efficiency in guiding you to useful information
( ) The system's ability to continue searching on its own to find useful information

Please answer two final questions on the subjects you search in library computer systems.

13. How would you rate your familiarity with the subject you searched today?
( ) Very familiar
( ) Somewhat familiar
( ) Somewhat unfamiliar
( ) Very unfamiliar

14. How interested are you in seeing the capabilities of ASTUTE extended to topics besides [Engineering and Computer Science (at UM-Dearborn)] or [American History (at Earlham)]?
( ) Very interested [prompt with question 15]
( ) Somewhat interested
( ) Somewhat uninterested
( ) Very uninterested

15. Please type in what topics you would like added to ASTUTE.

Appendix D.
Cover Letter and Consent Form Completed by Library Staff

[Date]

[Name and address of
prospective library staff participant]

Dear [prospective library staff participant]:

You are invited to participate in a study to test a new subject access design to online catalogs. This study involves searching the ASTUTE online catalog that has been installed in the reference area of [Mardigian or Lilly] Library since [date] and answering questions about the search that ASTUTE and I ask you. If you wish to participate, please fill out and sign the enclosed form, and return it to me in the enclosed envelope by [date]. At the same time as you are completing the form, [name of library liaison] will contact you about setting up a one-hour appointment to search ASTUTE in [date]. ([Name of library liaison] is helping me schedule the test to save me from making several than trips to [Mardigian or Lilly] Library.)

Please bring three to five subject queries in the area of [computer science or engineering; American history] to search. These queries could have come from library patrons in the course of your reference work or may be subject queries in which you have a professional o personal interest. I will not reveal your identity in reports generated by this study but I may discuss your queries and ASTUTE's ability to retrieve relevant titles for them.

Thank you for your assistance. If you have any questions about the study, please feel free to contact me by phone at 313-763-3581 or by electronic mail (karen.drabenstott@umich.edu).

Sincerely,

Karen M. Drabenstott
Associate Professor

Enclosure (1)

SCHOOL OF INFORMATION AND LIBRARY STUDIES
NEW SUBJECT ACCESS DESIGN STUDY
Consent Form

Please print:

Name:

Title:

Mailing address:

Office phone:

Electronic mail address:

Do you prefer to be contacted by
electronic mail? (Circle one:) Yes No Don't care

Please read the following and sign:

I have read the enclosed information (cover letter) about this study and agree to participate.

I understand that I will not be identified in reports generated from this study. I understand that my query and ASTUTE's ability to retrieve relevant titles for it may be discussed in such reports.

I understand that my participation is voluntary, and that I may withdraw at any time.

Signed ______________________________

Date ________________________________

Please return this form to:

Karen M. Drabenstott
University of Michigan
School of Information and Library Studies
304 West Engineering Building
550 East University Avenue
Ann Arbor, MI 48109–1092

Appendix E. Post-search Questionnaire Administered to Library Staff

Questionnaire: # _____
Library: __UM-D EC__ First system: __Blue Pinstripe__
Date: _______ Time: _________ Interviewer: __KD__

(Note that more space was left in the actual interview form for participants' answers to open-ended questions to enable the interviewer to write down participant responses.)

A. Introduction

Hello. I'm ________________ from the School of Information and Library Studies at the University of Michigan. We are conducting a study to find out how well the ASTUTE experimental online catalog helps library patrons and staff find library material on the topics that interest them. ASTUTE is actually 2 different online catalogs. The 2 ASTUTEs contain the same titles but they have different subject searching capabilities. As you search, take note of the differences between the two ASTUTEs so that we can talk about these differences later.

We'll take 1 topic at a time. When you start searching ASTUTE, the system will ask you some questions about yourself and your previous use of the system. Then it will ask you to enter your subject query in 1 of the 2 ASTUTE online catalogs. As you retrieve titles, ASTUTE will ask you to assess their relevance. When you are done searching and make some moves to start over, ASTUTE will switch to the second ASTUTE catalog and search for the original subject query you entered. After you finish searching, ASTUTE will ask you some questions about your searches and search results. Then I will ask you additional questions about your experiences searching the 2 ASTUTEs.

Let's get started. Please go ahead and start searching ASTUTE for 1 of the 3 topics you brought with you.

Time: ________________

(Respondent searches for 8–13 minutes)

B. Questions for Search Topic 1.

1. What was the impetus for the subject you searched in ASTUTE?

(Probe: troublesome patron query, developing a bibliography, responsibility for collection development, frequent patron request)

2. Did you have any difficulties searching for your subject in the first (Blue or Pinstripe) system?

(Probe: Give examples.)

3. What improvements could be made to the first (Blue or Pinstripe) system?

4. Did you have any difficulties searching for your subject in the second (Blue or Pinstripe) system?

(Probe: Give examples.)

5. What improvements could be made to the second (Blue or Pinstripe) system?

6. What differences did you notice between the ways in which the 2 systems responded to the subjects you typed in?

a. Probe for examples.

b. Probe: What do you think about this?

c. Probe: Can you tell me why?

d. What bearing does this have on your system preference?

7. In what ways did the (Blue or Pinstripe) system give you new ideas for searching for subjects?

a. Probe for examples.

b. Probe: What do you think about this?

c. Probe: Can you tell me why?

d. What bearing does this have on your system preference?

8. What differences did you notice between the titles the 2 (Blue and Pinstripe) systems found for your subject?

a. Probe for amount of titles, proportion of relevant titles, proportion of nonrelevant titles.

b. What bearing does this have on your system preference?

Let's move onto the next search. Please go ahead and start searching the (Blue or Pinstripe) system for the second of the 3 topics you brought with you.

Time: ________________
(Respondent searches for 8-13 minutes)

C. Questions for Search Topic 2.

(Repeat questions B.1–9 above.)

Let's move onto the last search. Please go ahead and start searching the (Blue or Pinstripe) system for the third topic you brought with you.

Time: ________________
(Respondent searches for 8–13 minutes)

D. Questions for Search Topic 3.

(Repeat questions B.1–9 above.)

E. Summary Questions

Imagine that we could transform your library's online catalog into a new catalog using the best features of this experimental online catalog.

10. What do you like most about the Blue or Pinstripe system?

(Probe: which system, give examples of features.)

11. What do you like least about the Blue or Pinstripe systems?

(Probe: which system, give examples of features, feeling of control, etc.)

12. Overall, which system do you prefer?

a. Can you tell me why?

13. Are there any additional comments you would like to make concerning your use of ASTUTE?

THANK YOU for participating in this study.
Your comments will help us make recommendations
for improving subject access
in online catalogs.

Appendix F.
ASTUTE Transaction Log Records

F.1 Level 1 Transaction Log Records

ASTUTE assigned participants unique identification numbers in the order in which they used the system. Level 1 transaction log records contained these unique identification numbers, start and end times for the user's interaction with ASTUTE, and answers to the pre- and post-search questionnaires. User identification numbers were the same across all four levels of transaction log records, and, thus, served as a link between records at different levels. ASTUTE wrote one level 1 transaction log record per system user. When users walked away from their search before finishing, the system stopped writing values to level 1 record fields and left them blank. Table F1 lists fields of level 1 transaction log records.

Table F1. Level 1 Transaction Log Record Fields

Field description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Date	L1_Date	character	8
Search experiment start time	L1_St_Time	character	8
Search experiment end time	L1_En_Time	character	8
First system searched	L1_Sys_Ord	character	1
Answer to pre-search question 1	L1_Q1	character	1
Answer to pre-search question 2	L1_Q2	character	1
Answer to pre-search question 3	L1_Q3	character	2
Time immediately prior to user's answer to question 4	L1_Q4_Time	character	8
Answer to post-search question 4	L1_Q4	character	1
Answer to post-search question 5	L1_Q5	character	1
Answer to post-search question 6	L1_Q6	character	1

Answer to post-search question 7	L1_Q7	character	1
Answer to post-search question 8	L1_Q8	character	1
Answer to post-search question 9	L1_Q9	character	1
Answer to post-search question 10	L1_Q10	character	1
Answer to post-search question 11	L1_Q11	character	1
Answer to post-search question 12, first-listed category	L1_Q12_1	character	1
Answer to post-search question 12, second-listed category	L1_Q12_2	character	1
Answer to post-search question 12, third-listed category	L1_Q12_3	character	1
Answer to post-search question 12, fourth-listed category	L1_Q12_4	character	1
Answer to post-search question 12, fifth-listed category	L1_Q12_5	character	1
Answer to post-search question 12, sixth-listed category	L1_Q12_6	character	1
Answer to post-search question 13	L1_Q13	character	1
Answer to post-search question 14	L1_Q14	character	1

Table F2 shows a level 1 transaction log record for ASTUTE user #6 at Earlham College. Answers to pre-search questions 1–3 tell us this user is using ASTUTE for the first time, uses other computer systems on a weekly basis, and has a major field of study in the humanities.

Table F2. Complete Level 1 Transaction Log Record

Field abbreviation	Answer code	Answer category
U_Id_No	6	(User identification number)
L1_Date	2/23/93	(Date)
L1_St_Time	21:02:32	(Search experiment start time)
L1_En_Time	21:13:37	(Search experiment end time)
L1_Sys_Ord	P	Pinstripe (first system searched)
L1_Q1	1	"Never before"
L1_Q2	3	"Weekly"
L1_Q3	7	"Other humanities fields"
L1_Q4_Time	21:11:30	Time immediately prior to user's answer to question 4
L1_Q4	2	"About the same number of useful titles in both systems"
L1_Q5	3	"No difference"

L1_Q6	3	"No difference"
L1_Q7	3	"Both are easy"
L1_Q8	3	"Both have useful instructions"
L1_Q9	3	"Both are clear"
L1_Q10	3	"Both are helpful"
L1_Q11	4	"Neither is efficient"
L1_Q12_1	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_2	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_3	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_4	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_5	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_6	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q13	4	"Very unfamiliar"
L1_Q14	2	"Somewhat interested"

Post-search questionnaire responses indicate that this user felt that there was no difference between the performances of the Blue and Pinstripe Systems. The user felt the systems were easy to use and understand but neither was especially efficient about retrieving useful information. Although he rated all capabilities listed in question 12 as being important, he felt that the ease of giving instructions to the system, system clarity, and system efficiency in guiding him to useful information were the most important aspects that helped him search library computer systems. The user was very unfamiliar with the subject he searched in the Blue and Pinstripe Systems and was somewhat interested in seeing the systems' capabilities extended to topics besides American history.

F.2 Level 2 Transaction Log Records

Level 2 transaction log records documented the queries users entered into ASTUTE and the subject searching approaches the system enlisted to retrieve subject headings or titles. ASTUTE wrote as many level 2 records as necessary to document user queries and subject searching approaches. The user identification number field of level 2 records linked them to level 1, 3, and 4 records. The series number field of level 2 records linked level 2 records to one another. Table F3 lists fields in level 2 transaction log records.

Table F3. Level 2 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Level 2 series number	L2_Series	numeric	3
System user is searching	L2_Sys	character	1
Search approach in use	L2_M_Approach	character	1
User-induced approach	L2_Appr	character	5
Time level 2 record is written	L2_Time	character	8
Match made or not made using this approach	L2_Results	numeric	1
Number of titles retrieved	L2_Titles	character	5
System- or user-induced method of continuing the search	L2_Contype	character	4
User-entered query	L2_Query	memo	N/A

Level 2 records contained ten fields. User identification and series number fields linked level 2 records to each other and transaction records for other levels. Level 2 records documented the system the user is searching with a one-character code, i.e., B (Blue System) or P (Pinstripe System). The L2_Query field stored user-entered queries or user-selected terms in a variable-length memo field that contained as many characters as the queries or terms. The system wrote a one-character code in the L2_M_Approach field to document the subject searching approach the Blue or Pinstripe System used to produce retrievals for the user query. Approaches and codes are given in table F4.

Table F4. Codes Used in L2_M_Approach Fields

Codes	Subject search approach
1	Exact search
2	Alphabetical search
3	Keyword-in-main-heading search
4	Keyword-in-subdivided-heading search
5	Title-keyword search for queries exceeding one word
6	Keyword in subject heading fields search
7	Keyword-in-record search
8	Title-keyword search for one-word queries
9	Keyword-in-record search for personal names

A	Alphabetical search for personal names
0	(System switch to Blue or Pinstripe)

The system wrote a 1 (one) for success or 0 (zero) for failure to the L2_Results field to document whether the search approach coded in the L2_M_Approach field was successful. If the system was successful, it gave users options to further the search. For example, in response to a successful exact search, the system could give users several methods of furthering the search, e.g., displaying broader terms, displaying narrower terms, displaying scope notes, displaying subdivided forms of subject headings. The system wrote a code in the L2_Appr field to document the method that users selected to further the search. This field could also contain codes to indicate the type of subject search that the system was executing or the system's response to a user action to start over. The latter ran the gamut of subject searching capabilities implemented in the experimental online catalog. The former were limited to codes for "swit" (i.e., switch) to indicate that the Blue System was switching to the Pinstripe System or vice versa and to "quit" to indicate that the system was initiating the post-search questionnaire. Table F5 lists the three types of codes written to the L2_Appr field of level 2 records.

Table F5. Codes Used in L2_Appr Fields

Codes	Code Descriptions
Subject search codes:	
ex	System is executing the exact approach
alph	System is executing the alphabetical approach
kwmh	System is executing the keyword-in-main-heading search
kwt	System is executing the title-keyword search
kwsf	System is executing the keyword in subject heading fields search
kwr	System is executing the keyword-in-record search
pkwr	System is executing the keyword-in-record search for personal-name subject queries
palp	System is executing the alphabetical approach for personal-name subject queries
Codes for furthering searches:	
xref	Within the exact or alphabetical approach, the system is displaying a see reference
notes	Within the exact approach, the system is displaying a scope note
titles	System is displaying retrieved bibliographic records
nt	Within the exact approach, the system is displaying narrower terms
bt	Within the exact approach, the system is displaying broader terms

subj	Within the exact approach, the system is displaying subject headings subdivided by topical subdivisions
geog	Within the exact approach, the system is displaying subject headings subdivided by geographical subdivisions
chron	Within the exact approach, the system is displaying subject headings subdivided by chronological subdivisions
mat	Within the exact approach, the system is displaying subject headings subdivided by form subdivisions
Housekeeping codes:	
swit	System is switching between Blue and Pinstripe Systems
quit	System is initiating post-search questionnaire

The L2_Contype field documented the actions taken by users to continue the search. Table F6 lists and describes the codes the system wrote to this field.

Table F6. Codes Used in L2_Contype Fields

Codes	Subject search approach
next	Blue System automatically goes to next approach on search trees
more	User makes an action to further the approach. Examples: selecting headings, browsing narrower terms, displaying scope notes
brow	User browses a list but does not choose a listed subject heading
expd	User chooses to expand the search which induces the Blue System to go to next approach on search trees
prev	User goes to a previous screen in exact approach
swit	System switches to Blue or Pinstripe System
quit	User makes move to start over and system initiates post-search questionnaire
nfnd	System is unable to find user-entered query in keyword dictionary

Level 2 records also contained fields for the time of individual transactions. Numbers coded in the L2_Titles field designated the number of bibliographic records users retrieved for the queries they entered or subject headings they selected.

Table F7 shows the series of level 2 records linked to the level 1 record in Table F2. The user entered the query "richmond, indiana" into the Pinstripe System. The system wrote only one level 2 record for the search in the Pinstripe System. When the user made a move to start over, the system switched to the Blue System and executed a search for the original query, i.e., "richmond, indiana." It wrote eleven level 2 records for user actions and system responses in the Blue System. In the level 2 records in Table F6, L2_Query fields bearing user-entered or user-selected terms are given under the nine other fields of these records.

Table F7. Complete Series of Level 2 Transaction Log Records

Id	Series	Sys	M_Ap	Appr	Cont	Time	Result	Titles
6	1	P	7	titles	next	21:03:05	1	16
richmond, indiana								
6	2	B	0	swit	swit	21:05:37	0	0
richmond, indiana								
6	3	B	1	ex	next	21:05:48	0	0
richmond, indiana								
6	4	B	2	alph	next	21:05:48	0	0
richmond, indiana								
6	5	B	3	kwmh	next	21:05:48	0	0
richmond, indiana								
6	6	B	4	kwsh	next	21:06:26	1	0
richmond, indiana								
6	7	B	4	kwsh	more	21:06:26	1	0
Historic buildings — Indiana — Richmond								
6	8	B	4	titles	more	21:06:40	1	1
Historic buildings — Indiana — Richmond								
6	9	B	4	kwsh	next	21:07:26	1	0
Historic buildings — Indiana — Richmond								
6	10	B	4	kwsh	expd	21:07:26	1	0
richmond, indiana								
6	11	B	5	titles	next	21:07:37	1	13
richmond, indiana								
6	12	B	5	kwt	quit	21:11:29	1	0
richmond, indiana								

Key to table headings:

Id = U_Id_No
Series = L2_Series
Sys = L2_System
Appr = L2_Approach
Cont = L2_Contype
Time = L2_Time

Result = L2_Result
M_Ap = L2_M_Approach
Titles = L2_Titles

Level 2 transaction log records demonstrate that the user searched the Pinstripe System first with the query "richmond, indiana." The system wrote code 7 in the L2_M_Approach field to document that the Pinstripe System randomly chose the keyword-in-record search for this query. The system retrieved sixteen bibliographic records (L2_Titles field). To determine whether the user displayed any of these records, we need to inspect level 3 records. The user made an action to terminate the search (L2_Contype field). In response, ASTUTE switched to the Blue Test System (L2_System field).

The Blue System executed the exact approach, alphabetical approach, and keyword-in-main-heading search (L2_Series, L2_M_Approach, and L2_Result fields) without success. The sixth record in the level 2 record series documented that the Blue System executed a keyword-in-subdivided-heading search (L2_M_Approach). The "1" in the L2_Result field indicates that the Blue System was successful in retrieving subject heading(s). The user selected a listed heading (L2_Contype). The heading is listed under coded data in the seventh record in the level 2 record series, i.e., "Historic buildings — Indiana — Richmond." This subject heading was connected with one title (L2_Titles field). The tenth record in the level 2 record series indicates that the user wanted the system to automatically expand the search (L2_Contype). This means the user pursued a system prompt that induced the Blue System proceed with the next approach on search trees to produce retrievals. The system resumed with a title keyword search (L2_M_Approach) and was successful in retrieving thirteen titles (L2_Titles field). The contents of level 3 records are needed to determine whether the user displayed one or more of the thirteen records. When the user made a move to terminate the search, the system stopped writing level 2 records and resumed with the post-search questionnaire. (Level 1 records documented questionnaire responses.)

F.3 Level 3 Transaction Log Records

Level 3 transaction log records contained fields to document bibliographic records that users displayed and rated. They also contained fields to enable the researchers to link level 3 records to level 1 and 2 records. Table F8 lists fields in level 3 transaction log records.

Table F8. Level 3 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Level 2 series number	L2_Series	numeric	3
Level 3 series number	L3_Series	numeric	4
Time level 3 record is written	L3_Time	character	8
Bibliographic record control number	Bib_Series	numeric	5
User relevance assessment	L3_Relevance	numeric	1

Serial numbers of the bibliographic records that users displayed were given in the Bib_Series field. User relevance assessments were assigned to the L3_Relevance field. The meaning of codes 1–3 in L3_Relevance field was: (1) "not useful," (2) "possibly useful," and (3) "very useful." Of the fields bearing numbers linking level 3 records to corresponding level 1 and 2 records, unique user identification numbers were given in U_Id_No fields. Level 3 records contained level 2 series numbers (L2_Series) to enable the researchers to determine the particular user activity that led to the retrieval of bibliographic records. Time stamps were given in L3_Time fields.

Table F9 shows level 3 transaction log records written for a user at Earlham College. Level 1 and 2 records for this user were shown in Tables F2 and F7. Level 2 records demonstrate that the user retrieved bibliographic records in level 2 series records 1, 8, and 11 (Table F7). Thus, the user had the opportunity to display records as a result of searching activity connected with level 2 series records 1, 8, and 11.

Table F9. Complete Series of Level 3 Transaction Log Records

U_Id_No	L2_Series	L3_Series	L3_Time	Bib_Series	L3_Relevance
6	1	1	21:03:05	10883	2
6	1	2	21:03:20	10821	1
6	1	3	21:03:32	10819	1
6	1	4	21:03:41	10811	1
6	1	5	21:03:52	10868	1
6	1	6	21:04:02	10739	1
6	1	7	21:04:07	10747	2
6	1	8	21:04:20	10869	1
6	1	9	21:04:26	10804	2
6	1	10	21:04:38	10888	1
6	1	11	21:04:44	10508	1

6	1	12	21:04:51	10858	1
6	1	13	21:04:59	10871	2
6	1	14	21:05:07	10746	1
6	1	15	21:05:14	10864	1
6	1	16	21:05:29	10887	1
6	8	1	21:06:47	10821	1
6	11	1	21:07:38	10883	2
6	11	2	21:07:52	10821	1
6	11	3	21:07:56	10811	1
6	11	4	21:08:00	10868	1
6	11	5	21:08:11	10739	1
6	11	6	21:08:15	10747	2
6	11	7	21:08:25	10804	1
6	11	8	21:08:31	10888	1
6	11	9	21:08:35	10858	1
6	11	10	21:09:37	10871	2
6	11	11	21:09:43	10746	2
6	11	12	21:10:44	10864	1
6	11	13	21:10:56	10887	1

In the Comparison Search Experiment conducted by the user at Earlham College, the system wrote thirty level 3 records for displayed and rated bibliographic records. The user displayed all sixteen records that the Pinstripe System retrieved in a keyword-in-record search. The user took two minutes and twenty-four seconds to display and rate the usefulness of the sixteen records. He gave four records "possibly useful" ratings and twelve records "not useful" ratings. The user retrieved fourteen bibliographic records in the Blue System. The user gave a "not useful" rating to the one record retrieved through the keyword-in-subdivided-heading search (L2_Series field = 8). Of the thirteen records retrieved through the title-keyword search (L2_Series fields = 11), the user rated four records "possibly useful" and nine records "not useful." It took the user three minutes and eighteen seconds to review these thirteen records. The fourteen records displayed in Blue-system searches were displayed in Pinstripe-system searches. User relevance assessments for two of the fourteen titles displayed in searches on both systems varied (Bib_series #10804 and #10746). (Chapter 8 described bibliographic records that users displayed most than once and the resolution of conflicting relevance assessments.)

F.4 Level 4 Transaction Log Records

Level 4 records were written under two conditions: (1) users selected the "very interested" category in response to post-search question #14 on their interest in seeing ASTUTE's capabilities extended to other topics, and (2) users typed something into the pop-up box asking them for suggested topics. Level 4 records contained only two fields: (1) the unique user identification number to link level 4 records to level 1, 2, and 3 records (U_Id_No), and (2) the terms users typed into the pop-up box (L4_Topics). Table F10 lists fields in level 4 transaction log records.

Table F10. Level 4 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
User-entered suggested topic(s)	L4_Topics	memo	N/A

User identification numbers were limited to fixed-length fields. User-entered suggested topics were saved in a variable-length memo field. The system did not write a level 4 record for the user at Earlham because he did not select the "very interested" category in response to post-search question #14. Several ASTUTE users at Earlham College did select this category and their suggested topics are given in Table F11.

Table F11. Examples of Level 4 Transaction Log Records

U_Id_No	L4_Topics
31	i don't know
35	religion/philosophy
92	2
102	Child abuse
113	AFRICAN AMERICAN WOMEN WRITERS IN HARLEM RENAISSANCE
146	life
151	Jump Rope Through the Ages
164	supreme court decisions
193	po;itics
201	women in reconstruction; civil war

Although subsequent chapters describe user-suggested topics, it is interesting to see the wide range of comments. User-suggested topics could be very broad (e.g., "life") or very specific (e.g., "Jump Rope Through the Ages"). They could also be unintelligible (e.g., "2"), contain typographical errors (e.g., "po;itics"), or say what was on the user's mind (e.g., "i don't know").